

**MODELOS ESTADÍSTICOS TIPO ARIMA PARA EL PRONÓSTICO DE
INCAUTACIONES DE DROGAS ILEGALES POR PARTE DE LA ARMADA
NACIONAL DE COLOMBIA**



**BREYNER JIMÉNEZ NAVIA
CARLOS ADOLFO MORENO BUITRAGO**

**FUNDACIÓN UNIVERSITARIA LOS LIBERTADORES
DEPARTAMENTO DE CIENCIAS BÁSICAS
ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA
BOGOTÁ, D. C. 2017**

**MODELOS ESTADÍSTICOS TIPO ARIMA PARA EL PRONÓSTICO DE
INCAUTACIONES DE DROGAS ILEGALES POR PARTE DE LA ARMADA
NACIONAL DE COLOMBIA**

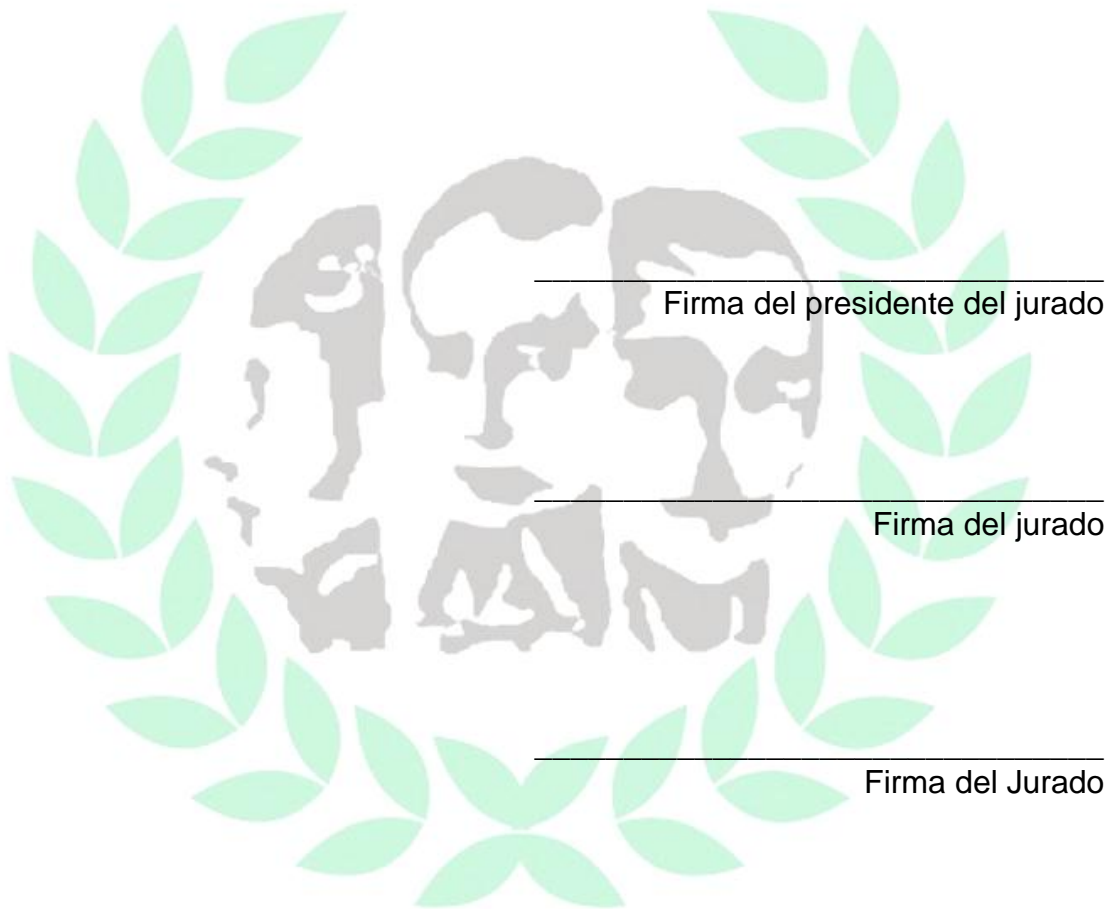
**BREYNER JIMÉNEZ NAVIA
CARLOS ADOLFO MORENO BUITRAGO**

(Trabajo de grado para optar al título de Especialista en Estadística Aplicada)

Director: Juan Camilo Santana (MSc en Economía y Estadística)

**FUNDACIÓN UNIVERSITARIA LOS LIBERTADORES
DEPARTAMENTO DE CIENCIAS BÁSICAS
ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA
BOGOTÁ, D. C. 2017**

Nota de aceptación



Bogotá, D.C., 15 de junio de 2017



Las directivas de la Fundación Universitaria los Libertadores, los jurados calificadores y el cuerpo docente no son responsables por los criterios e ideas expuestas en el presente documento. Estos corresponden únicamente a los autores

AGRADECIMIENTOS

A mis familiares, especialmente a mi madre, hermana, esposa e hijo, por la paciencia y el cariño con que me han acompañado durante todo el año en este proceso de aprendizaje; de igual forma a la Armada Nacional de Colombia que fue la que me impulsó y me llevó a adquirir esta nueva competencia en el campo profesional.

Breyner Jiménez Navia

A mis padres, quienes a su manera me han animado a buscar logros más altos cada día; a mis compañeros de trabajo en el colegio El Virrey José Solís, por el apoyo y motivación que me han brindado para poder llevar a cabo esta empresa; a mis compañeros de estudio, especialmente a Breyner Jiménez, por su compañerismo y motivación; a los docentes de la Especialización en Estadística Aplicada, por su ejemplo de compromiso y profesionalismo.

Carlos Adolfo Moreno Buitrago

CONTENIDO

	pág.
INTRODUCCIÓN	12
1. MARCO DE REFERENCIA	14
1.1 CULTIVOS ILÍCITOS	16
1.2 INCAUTACIONES	17
1.3 PRECIOS Y MERCADO	20
2. MARCO TEÓRICO	23
2.1 ANÁLISIS DE SERIES TEMPORALES	23
2.1.1 Tipos de análisis de series de tiempo	25
2.1.2 Procesos estocásticos	25
2.2 MODELOS ESTADÍSTICOS DE SERIES DE TIEMPO	26
2.2.1 Componentes inobservados de las series de tiempo	26
2.2.2 Modelo clásico de series de tiempo	26
2.2.3 Modelos ARIMA (enfoque Box Jenkins)	28
2.2.4 Evaluación de la capacidad predictiva en modelos ARIMA	29
3. MARCO METODOLÓGICO	30
3.1 RECURSOS	31
3.2 IMPACTOS	31
4. ANÁLISIS Y RESULTADOS	32
4.1 SERIE INCAUTACIONES DE COCAÍNA 2012-2016	32
4.2 PRUEBAS DE LOS MODELOS ARIMA Y ARIMAX INICIALES	35
4.3 PRUEBAS DE LOS MODELOS ARIMAX SIN ANÁLISIS DE INTERVENCIÓN	36
4.3.1 Prueba del modelo ARIMAX (1.1.0) con 01 regresor	37
4.3.2 Prueba del modelo ARIMAX (6.1.2) con 01 regresor	39
4.3.3 Prueba del modelo ARIMAX (2.1.4) con 02 regresores	41
4.4 PRUEBAS DE LOS MODELOS ARIMAX CON ANÁLISIS DE INTERVENCIÓN	43
4.4.1 Prueba del modelo ARIMAX(7.1.1) con análisis de intervención de 05 dummies	44
4.4.2 Prueba del modelo ARIMAX (7.1.1) con análisis de intervención de 06 dummies tipo puntual y una de cambio de nivel (LS)	47
4.4.3 Pruebas de capacidad predictiva de los modelos ARIMAX(7.1.1) con intervención de 07 dummies, con 05 dummies y con 07 dummies (una de cambio de nivel - LS)	49
5. CONCLUSIONES Y RECOMENDACIONES	51
REFERENCIAS	52

LISTA DE FIGURAS

	pág.
Figura 1. Incautaciones de clorhidrato de cocaína en escenarios marítimo y fluvial por parte de la Armada Nacional en 2015	17



LISTA DE TABLAS

	pág.
Tabla 1. Variables sobre narcotráfico de interés para la Armada Nacional	15
Tabla 2. Precios promedio de la hoja de coca y sus derivados, 2008-2013	21
Tabla 3. Serie de tiempo de número de reservas de pasajeros internacionales en la aerolínea PanAm en los Estados Unidos	23
Tabla 4. Serie de tiempo de incautaciones de cocaína en kg realizadas por la Armada Nacional	32
Tabla 5. Comparativo de las pruebas de los modelos ARIMA Y ARIMAX iniciales	36
Tabla 6. Comparativo de las pruebas de los modelos ARIMAX sin análisis de intervención	36
Tabla 7. Comparativo de las pruebas de los modelos ARIMAX con análisis de intervención	43
Tabla 8. Diferencias mensuales en las predicciones de los tres modelos en relación con los datos reales	50
Tabla 9. Mejor ajuste (MAPE) de los modelos ARIMAX (7.1.1) con 07 dummies, el ARIMAX (7.1.1) con 05 dummies y el ARIMAX (7.1.1) con 06 dummies puntuales y una dummie LS	50

LISTA DE GRÁFICAS

	pág.
Gráfica 1. Serie de tiempo cultivos de coca	16
Gráfica 2. Cultivos de coca vs. esfuerzos para su control	17
Gráfica 3. Hoja de coca histórica.	18
Gráfica 4. Pasta básica/base de cocaína histórico.	18
Gráfica 5. Cocaína histórica.	19
Gráfica 6. Incautación de sustancias químicas sólidas 2001-2012 (kg).	19
Gráfica 7. Incautaciones de sustancias químicas líquidas según año	20
Gráfica 8. Precio internacional de clorhidrato de cocaína por kg	22
Gráfica 9. Serie de tiempo del consumo mensual de gasolina en España entre enero de 1966 y agosto de 1977	27
Gráfica 10. Serie original del consumo mensual de gasolina en España (enero de 1966 hasta agosto de 1977) y su descomposición en componentes aditivos	28
Gráfica 11. Serie de tiempo de incautaciones mensuales de kilogramos de cocaína realizadas por la Armada Nacional 2012-2016	32
Gráfica 12. Serie de tiempo de incautaciones de cocaína (kg) y correlogramas	33
Gráfica 13. Serie de tiempo transformada BoxCox de cocaína (kg)	34
Gráfica 14. Serie de tiempo de incautaciones de cocaína transformada con BoxCox y una (01) diferencia	35
Gráfica 15. Prueba del modelo ARIMAX (1.1.0) con 01 regresor	37
Gráfica 16. Prueba de homocedasticidad del modelo ARIMAX (1.1.0) con 01 regresor	38
Gráfica 17. Serie de tiempo original y del modelo ARIMAX (1.1.0) con 01 regresor y pronóstico	39
Gráfica 18. Prueba del modelo ARIMAX (6.1.2) con 01 regresor	39
Gráfica 19. Prueba de homocedasticidad del modelo ARIMAX (1.1.0) con 01 regresor	40
Gráfica 20. Serie de tiempo original, del modelo ARIMAX (6.1.2) con 01 regresor y pronóstico	41
Gráfica 21. Prueba del modelo ARIMAX (2.1.4) con 02 regresores	41
Gráfica 22. Prueba de homocedasticidad del modelo ARIMAX (2.1.4) con 02 regresores	42
Gráfica 23. Serie de tiempo original, del modelo ARIMAX (2.1.4) con 02 regresores y pronóstico	43
Gráfica 24. Serie de tiempo transformada y algunos valores atípicos	44
Gráfica 25. Prueba del modelo ARIMAX (7.1.1) con análisis de intervención de 05 dummies	45
Gráfica 26. Prueba de homocedasticidad del modelo ARIMAX (7.1.1) con 05 dummies de Intervención	46
Gráfica 27. Serie de tiempo original y del modelo ARIMAX (7.1.1) con 06 dummies de Intervención y pronóstico	46

Gráfica 28. Prueba del modelo ARIMAX (7.1.1) con análisis de intervención de 06 dummies de tipo aditivo y una de cambio de nivel.	47
Gráfica 29. Prueba de homocedasticidad del modelo ARIMAX (7.1.1) con análisis de intervención de 06 dummies de tipo aditivo y una de cambio de nivel	48
Gráfica 30. Serie de tiempo original y del modelo ARIMAX (7.1.1) con intervención de 06 dummies de tipo AO y una de tipo LS y pronóstico	49



RESUMEN

En este escrito se recoge el interés de la Armada Nacional de Colombia (ARC) por consolidar una metodología confiable, cuantificable y que se pueda nutrir con información estadística de las variables referidas al fenómeno del narcotráfico sobre las cuales se realizan las proyecciones de metas operacionales. De esta forma, se buscó desarrollar un conjunto de modelos estadísticos tipo ARIMA que permitan realizar las proyecciones de metas operacionales de incautaciones de cocaína por parte de la ARC y que posibiliten incorporar los registros históricos de resultados operacionales de años anteriores. Para tal fin, se realizó un proceso de análisis de series temporales bajo la metodología Box-Jenkins para modelos univariados tipo ARIMA. Después de probar veintiún (21) modelos diferentes, el último de los cuales incorporó variables regresoras relacionadas con las incautaciones de cocaína, se estableció que un modelo ARIMAX (7.1.1) con 05 dummies es el más conveniente para pronosticar las incautaciones de cocaína, debido a que los datos pronosticados con dicho modelo presentan menores diferencias frente a los datos reales.

ABSTRAC

In this paper, the National Navy of Colombia (ARC) is interested in consolidating a reliable, quantifiable methodology that can be fed with statistical information on the variables related to the phenomenon of drug trafficking, on which the operational targets are projected. In this way, we sought to develop a set of ARIMA-type statistical models that allow the projection of operational targets for cocaine seizures by the ARC and to make it possible to incorporate the historical records of operational results of previous years. For this purpose, a series of time series analysis was carried out under the Box-Jenkins methodology for ARIMA type univariate models. After testing twenty-one (21) different models, the last of which incorporated regressive variables related to cocaine seizures, it was established that an ARIMAX (7.1.1) model with 05 dummies is the most convenient to predict cocaine seizures, Due to the fact that the data predicted with this model present smaller differences compared to the real data.

Palabras claves: Narcotráfico, Metodología, Metas operacionales, Series de tiempo, Modelos ARIMA.

INTRODUCCIÓN

Aunque todos los fenómenos delictivos que ocurren en nuestro país impactan negativamente de una u otra forma a las comunidades, a la economía, al cumplimiento de los derechos humanos, al mantenimiento del orden constitucional y al bienestar general de la sociedad, no todos presentan las mismas dinámicas y expansión, y algunos de ellos se observan como generadores o perpetuadores de otros, tales como los cultivos ilícitos, que nutren al tráfico y fabricación de estupefacientes, y son los que directamente contribuyen al financiamiento y mantenimiento de grupos armados ilegales.

En razón al aumento de varios fenómenos delictivos en nuestro país, tales como el narcotráfico, la minería criminal, el contrabando, el tráfico de armas, municiones y explosivos entre otros; la Armada Nacional ha tenido que reevaluar constantemente sus metas operacionales, iniciando en su primera línea de mando, las de sus cuatro principales unidades operativas mayores o Fuerzas Navales (Fuerza Naval del Caribe, Fuerza Naval del Pacífico, Fuerza Naval del Oriente, Fuerza Naval del Sur), esto debido a que continuamente ha sucedido que las metas proyectadas para un (1) año por la Dirección de Operaciones Navales para estas unidades, resultan ser insuficientes o llegan a exceder sus capacidades operativas, por lo cual con frecuencia terminan siendo i) alcanzadas en la mitad del tiempo previsto o ii) imposibles de cumplir en el lapso de la vigencia.

El modelo actual de proyección de metas operacionales de la Dirección de Operaciones Navales presenta varias limitaciones en relación a los objetivos de hacer dichas proyecciones de manera más exacta, tomando en cuenta i) las capacidades operativas reales de las cuatro Fuerzas Navales, ii) las dinámicas de cambio de las variables asociadas a los fenómenos delictivos, iii) las asociaciones entre dichas variables y iv) la incorporación de los resultados operacionales de las Fuerzas Navales de años anteriores a partir de sus registros históricos.

Con el desarrollo de un nuevo modelo, de tipo estadístico, para la proyección de metas operacionales para las Fuerzas Navales de la Armada Nacional, se espera poder realizar proyecciones de metas más ajustadas a las capacidades operativas reales de las diferentes Fuerzas Navales, que sean suficientes para al menos una vigencia de evaluación (un año), y que se aumenten los niveles de cumplimiento de metas por parte las unidades operativas mayores, menores y tácticas.

La información contenida en este documento es propiedad de la Armada Nacional de Colombia y es privilegiada y confidencial y no debe ser reproducida, utilizada o divulgada al exterior de la organización o con propósitos externos a los de la organización sin la autorización respectiva.

Planteamiento del problema

La Armada Nacional requiere definir las metas operacionales con base en un método confiable, cuantificable y que se pueda nutrir con información estadística de las variables de interés con base en los datos históricos existentes.

Formulación del problema

Cuál es el modelo estadístico bajo el enfoque Box Jenkins mas apropiado para realizar las predicciones de incautaciones de clorhidrato de cocaína en el corto plazo para la Armada Nacional de Colombia.

Objetivo general

Elaborar un modelo estadístico confiable para proyectar las metas operacionales de incautaciones de clorhidrato de cocaína para la Armada Nacional.

Objetivos específicos

- Realizar el análisis de la serie de tiempo de incautaciones de clorhidrato de cocaína realizadas por la Armada Nacional en el periodo del año 2012 al 2016.
- Probar el modelo obtenido con pruebas de capacidad predictiva (back testing) para determinar el potencial del mismo.
- Pronosticar las metas mensuales de incautaciones de clorhidrato de cocaína para la Armada Nacional en la vigencia 2017.

1. MARCO DE REFERENCIA

El hecho de que la amenaza narcoterrorista de las últimas décadas haya impuesto la necesidad de emplear en toda su intensidad el poder militar de la Nación con el objeto de contenerla y contrarrestarla para, finalmente, llevarla a un estado que permita su neutralización definitiva, ha generado un impacto muy particular que se refleja en la composición actual de las Fuerzas Militares y su equipamiento, cuyo desarrollo se ha visto influenciado por la necesidad de tener que adaptarse a un conflicto asimétrico en el frente interno que demanda el empleo de medios y procedimientos estratégicos, operacionales y tácticos no convencionales, sin desentenderse de la responsabilidad constitucional de velar por la defensa nacional.

La Armada Nacional ejerce presencia y soberanía sobre el Mar Caribe, el Océano Pacífico, y todas las arterias fluviales navegables, con el propósito de mantener la integridad territorial, el orden constitucional y contribuir al desarrollo del poder marítimo-fluvial y a la protección de los intereses de la Nación.

Para controlar parte del territorio nacional que le ha sido asignado (territorio marítimo 928.660 km²: incluye mar caribe y océano pacífico; 2.900 km de costas: jurisdicción terrestre 40.920 km²; área de responsabilidad fluvial 24.725 km: aproximadamente 3.800 km son frontera fluvial) la institución cuenta con cuatro Fuerzas Navales, cada una de ellas ubicadas estratégicamente en los cuatro puntos cardinales de la geografía del país.

Estas unidades operativas mayores desarrollan operaciones ofensivas y de control, que están conformadas por varios componentes tácticos: naval, de infantería de marina, de guardacostas y aeronavales, las cuales cumplen su función principal de proporcionar defensa a la Nación ante amenazas, combatir a todo tipo de organizaciones que pretendan perturbar y/o desestabilizar el orden constitucional (delitos) o atentar contra la población civil por medio de acciones violentas o terroristas, para así garantizar la seguridad a los ciudadanos, sus recursos y sus derechos, con el objeto de crear condiciones que coadyuven a un desarrollo socioeconómico sostenible en Colombia.

El Ministerio de Defensa Nacional a través del Comando General de las Fuerzas Militares diseña un plan de guerra para contrarrestar todo tipo de amenazas para el país. Es por esto que la Armada Nacional diseña su plan de campaña operacional, el cual está inmerso en sus objetivos estratégicos, que son medidos a través de metas operacionales (variables que contrarrestan los fenómenos delictivos).

A través de la Jefatura de Operaciones Navales (Direcciones de Operaciones Navales - DIONA), que es un nivel estratégico operacional, se controlan todas las operaciones tácticas a nivel nacional y a su vez, por medio de esta, la Armada Nacional obtiene la información en tiempo real de cada uno de los resultados positivos que impactan a las organizaciones al margen de la ley.

El alto mando naval y la Direcciones de Operaciones Navales tienen la necesidad permanente de realizar seguimiento a las operaciones, para así dar un parte de tranquilidad al alto gobierno, de que se está realizando el trabajo y se están contrarrestando los fenómenos delictivos y cada una las variables que de manera directa financian y alimentan la criminalidad al interior del país.

La Dirección de Operaciones Navales es la unidad encargada de realizar el seguimiento permanente al cumplimiento de los objetivos estratégicos del plan de campaña a través del análisis del cumplimiento de las metas operacionales.

A continuación se relacionan las variables a las que la Armada Nacional le realiza seguimiento permanente:

Tabla 1. Variables sobre narcotráfico de interés para la Armada Nacional

Variable	Unidad de medida
Incautación de clorhidrato cocaína	kilogramos (kg)
Incautación de pasta base de coca	kilogramos (kg)
Incautación de marihuana	kilogramos (kg)
Destrucción de infraestructura para la producción de cocaína	unidades (uds)
Inmovilización de medios de transporte de cocaína	unidades (uds)
Erradicación de cultivos ilícitos	hectareas (ha)
Incautación insumos sólidos	kilogramos (kg)
Incautación insumos líquidos	galones (gal)
Incautación de combustible	galones (gal)

Fuente: Obtenido de Dirección de Operaciones Navales. Bogotá. 2017

Muchos de los fenómenos delictivos en Colombia que están asociados a las variables de interés para la Armada Nacional se generan o expanden, tanto cuantitativa como espacialmente, por la presencia de grupos al margen de la ley y el narcotráfico, tal como lo señalan Sánchez, Díaz y Formisano (2003). Por ejemplo, a partir de un análisis de la evolución del conflicto, la violencia y la actividad criminal en Colombia, dichos autores presentan unas hipótesis relacionadas:

- Existe una clara correlación entre la presencia espacial de esos grupos con la tasa de crimen violento, en particular de civiles no combatientes. El choque inicial de crimen violento genera una persistencia temporal y una difusión espacial que aumenta en forma permanente la tasa de homicidios, tanto en la unidad geográfica local como en la de los vecinos.
- El aumento del secuestro está claramente ligado a las mayores necesidades de financiación, por parte de los grupos armado ilegales.
- El narcotráfico es otro de los delitos ligados a la expansión de los grupos armados ilegales. En particular, el crecimiento de los cultivos ilícitos es el resultado de la propia expansión geográfica de esas organizaciones, en la medida en que les genera fuentes de financiación (Collier, 2001. Citado por Sánchez, Díaz y Formisano, Ibid., p. 33).

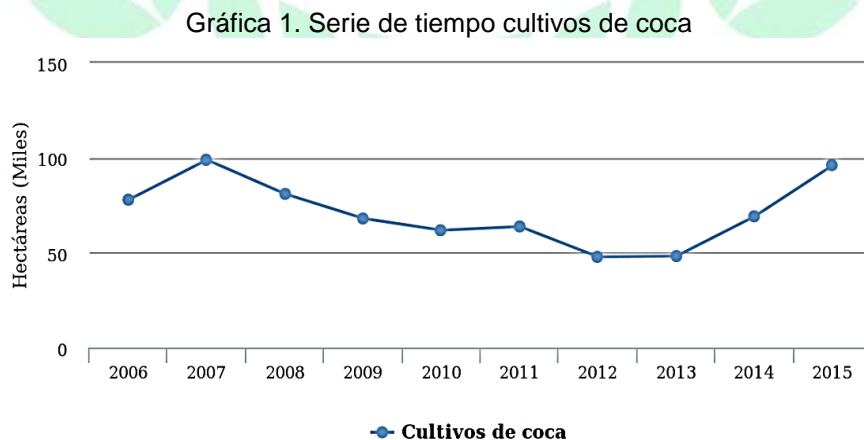
- La existencia de grupos armados ilegales puede facilitar el aumento de la delincuencia común y el aumento de otro tipo de delitos, en particular de delitos contra el patrimonio.
- Aunque las condiciones sociales como la desigualdad, la pobreza y la carencia de servicios sociales pueden ser el origen tanto del crimen violento como de los delitos contra el patrimonio (Fanzylber et al, 1998. Citados por Sánchez, Díaz y Formisano, Ibid., p. 33), y del nacimiento de grupos rebeldes, la dinámica del conflicto desencadena una dinámica propia del crimen violento y del delito común, en la que se pierde casi toda la relación con las condiciones sociales iniciales.

De esta forma, consideramos pertinente que un análisis de los resultados operacionales de las Fuerzas Navales de la Armada Nacional debe contemplar las dinámicas del conflicto interno colombiano, los fenómenos de asociación entre fenómenos delictivos e incluso las relaciones espaciales entre diferentes fenómenos delictivos, así como las dinámicas en el tiempo que han presentado los resultados operacionales de la Armada Nacional comparadas con las dinámicas del conflicto y los ciclos de violencia en Colombia.

Para el caso de las variables relacionadas con las drogas ilícitas, el Observatorio de Drogas de Colombia - ODC (2016) indica que en los últimos años ha habido una tendencia al aumento de varias de ellas.

1.1 CULTIVOS ILÍCITOS

En primer lugar, en cuanto a los cultivos de coca, después de haberse presentado una paulatina disminución de los mismos a partir del año 2007 y hasta el año 2013, se presenta ahora una tendencia a su aumento, llegando a 96.084,57 hs en 2015. La gráfica 1 muestra las cantidades de cultivos de coca sembrados en Colombia (en miles de hectáreas) según el año.

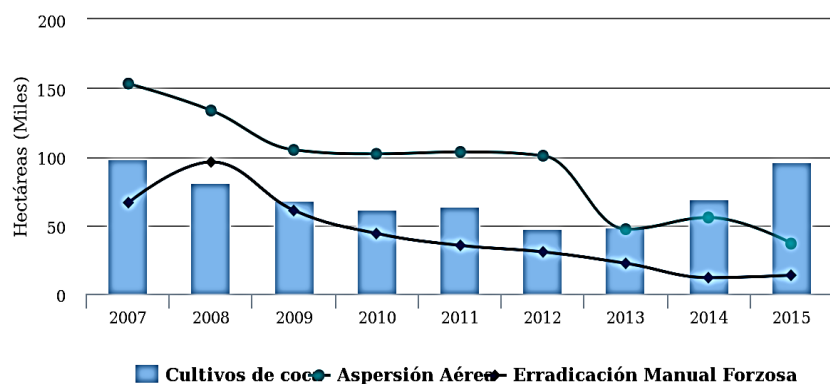


Fuente: ODC, 2016

De forma inconsistente, se encuentra que la erradicación de cultivos de coca ha tendido a disminuir, tanto para la estrategia de aspersión aérea (36.494,04 hs en

2015), como para la erradicación manual forzosa (13.473,46 hs). La gráfica 2 muestra estos comportamientos.

Gráfica 2. Cultivos de Coca vs. Esfuerzos para su control

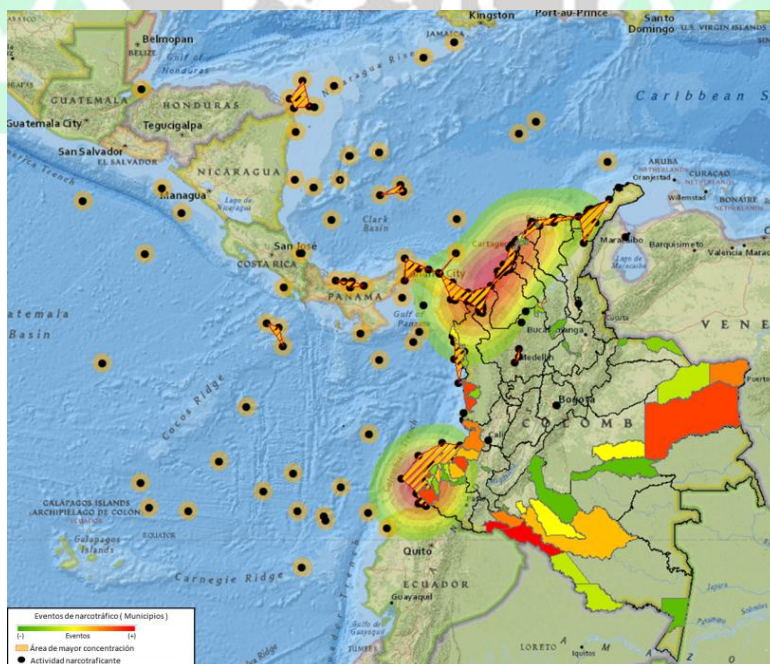


Fuente: ODC, 2016

1.2 INCAUTACIONES

Las incautaciones de clorhidrato de cocaína son una de las metas operacionales importantes para la Armada Nacional de Colombia, debido a que nuestro país ha sido por mucho tiempo uno de los principales productores de esta sustancia, y que en gran parte es destinada por los grupos de narcotráfico al exterior. La figura 1 muestra las zonas de Colombia y países vecinos donde la Armada Nacional realizó incautaciones de cocaína en 2015.

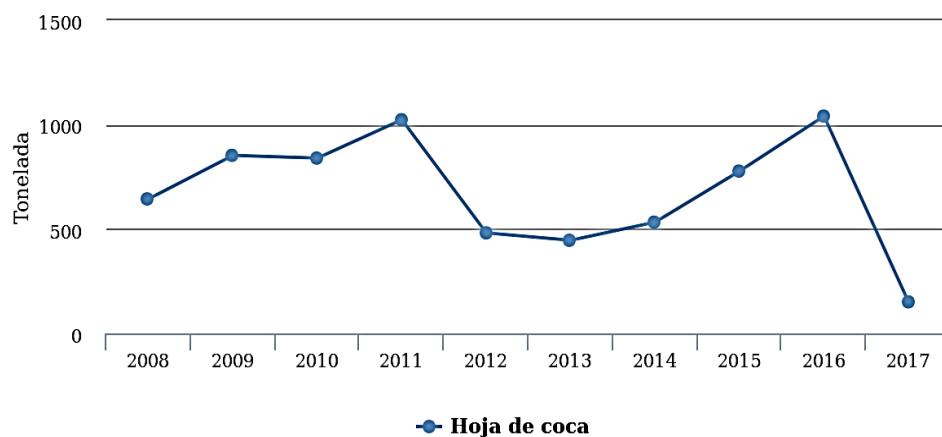
Figura 1. Incautaciones de clorhidrato de cocaína en escenarios marítimo y fluvial por parte de la Armada Nacional en 2015



Fuente: Dirección de Operaciones Navales, 2017

En cuanto a las incautaciones de drogas, el ODC presenta algunas mediciones importantes, como las toneladas de hoja de coca incautadas que se observan en la gráfica 3. En esta figura se observa que la tendencia ha sido al aumento (no se considera el presente año 2017 para esta medición).

Gráfica 3. Hoja de coca Histórico

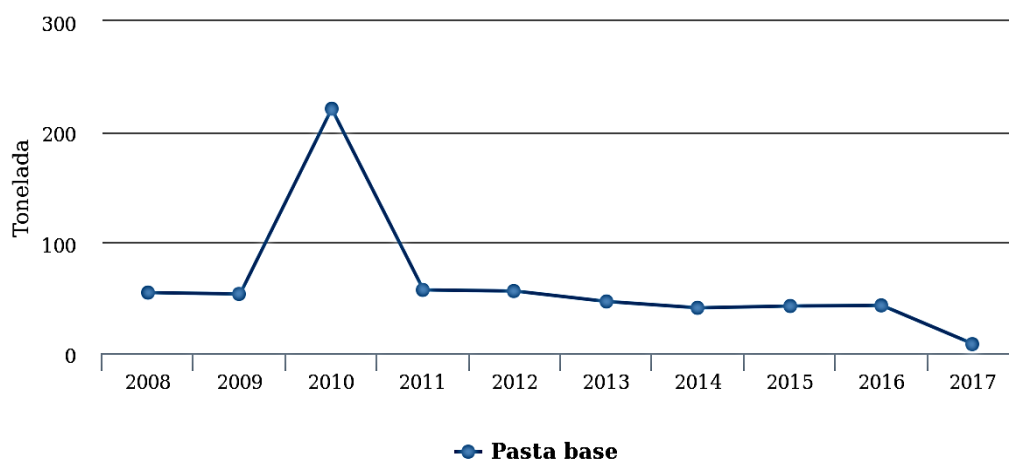


Fuente: ODC, 2016

De forma inconsistente, se encuentra que las incautaciones de pasta base de cocaína tienden a permanecer estables en el tiempo a partir de 2011 (en 2016 se incautaron 43,08 Tn). Esto lleva a pensar en varios factores que pueden afectar esta variable, como por ejemplo, la existencia de nuevas estrategias de los grupos armados y de narcotráfico para movilizar o almacenar las drogas, el aumento de rutas del narcotráfico al interior del país.

La gráfica 4 muestra la tendencia histórica de las incautaciones de pasta base de cocaína en Colombia.

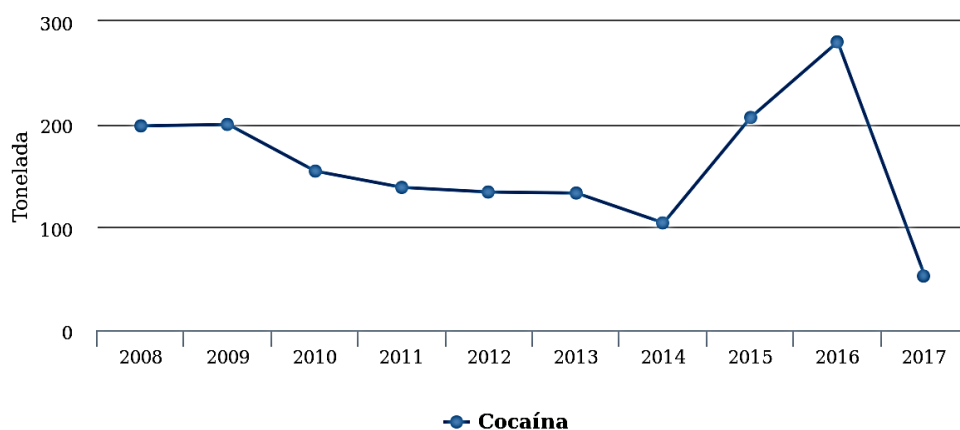
Gráfica 4. Pasta básica/base de cocaína Histórico



Fuente: ODC, 2016

La incautación de cocaína elaborada presenta un fuerte incremento a partir de 2014 (no se tiene en cuenta el presente año 2017), tal como se observa en la gráfica 5.

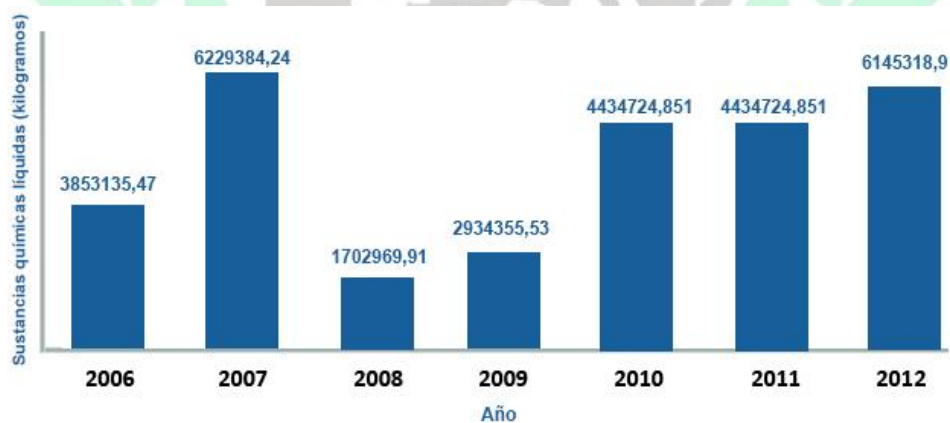
Gráfica 5. Cocaína histórico



Fuente: ODC, 2016

El ODC presenta también datos de incautaciones de sustancias químicas que se utilizan en los procesos de fabricación o extracción de las drogas ilícitas, ya sean estas de origen natural o sintético. Estas sustancias son bases, ácidos, solventes, oxidantes y reductores. La gráfica 6 muestra las cantidades de incautaciones de sustancias químicas sólidas en Colombia, dadas en kilogramos.

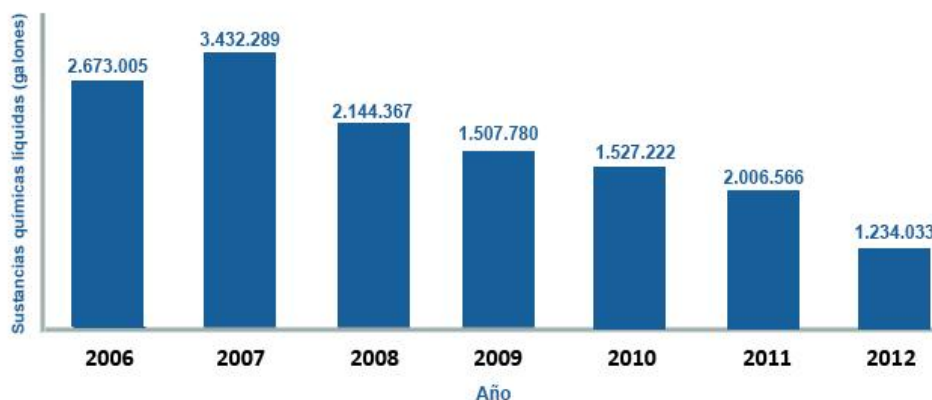
Gráfica 6. Incautaciones de sustancias químicas sólidas según año



Fuente: ODC, 2016

La gráfica 7 muestra las cantidades de incautaciones de sustancias químicas líquidas en Colombia, dadas en galones.

Gráfica 7. Incautaciones de sustancias químicas líquidas según año



Fuente: ODC, 2016

1.3 PRECIOS Y MERCADO

El mercado y los precios de los productos derivados de la producción y transformación de los cultivos ilícitos no dependen de las leyes de oferta y demanda, como ocurre con la mayoría de bienes y productos legales. En Colombia existen factores como el orden público (como la acción de las Fuerzas Militares y de Policía en una región, la acción de grupos armados ilegales, acción de grupos de delincuencia común, paros o movimientos civiles de comunidades), la oferta de hoja de coca fresca, el clima, acceso al transporte, cambios en los procesos productivos de las drogas, que llegan a incidir en alguna medida en el comportamiento de los precios en las regiones. De hecho, en algunas regiones los grupos armados u organizaciones de narcotráfico imponen reglas de juego que regulan el mercado en esas regiones (UNODC, 2014).

De acuerdo al ODC, hay varios eslabones en el mercado de las drogas ilícitas en Colombia: el primero es el productor, conformado por los campesinos que cultivan y colectan la hoja de coca, amapola o marihuana; este también puede encargarse (no necesariamente) de los procesos de extracción del alcaloide de la coca o recolección del opio (látex de la cápsula de la planta). El segundo eslabón son las personas encargadas del refinamiento, que buscan producir la pasta y base de cocaína; el tercer grupo lo conforman los encargados de la cristalización del clorhidrato de cocaína y la semi síntesis y síntesis de la morfina y la heroína (ODC, 2016).

Respecto al primer proceso de extracción del alcaloide de la coca, de acuerdo al informe *Monitoreo de Cultivos de Coca 2013* de la Oficina de las Naciones Unidas contra la Droga y el Delito (UNODC por sus siglas en inglés), en los años 2012 y 2013 se observó un aumento del número de campesinos productores que preferían vender la hoja de coca antes de procesarla para extraer la pasta básica de cocaína, lo que significó un 63% de los productores de hoja de coca en Colombia; de acuerdo a la UNODC, esto coincidió con el establecimiento de

nuevos laboratorios especializados para la extracción de pasta básica de cocaína y refinamiento a base de cocaína (UNODC, 2014).

Al analizar el funcionamiento del mercado global de la cocaína, que por cantidad es el producto ilegal más significativo en Colombia, se encuentra que los campesinos cultivadores en el primer eslabón son los más vulnerables, ya que solo reciben el 1,4% de los ingresos totales de la cocaína de todos los niveles del negocio, lo cual evidencia la desigualdad en este mercado.

De acuerdo al ODC (2014), el monitoreo de los precios contribuye a la mejor comprensión de la dinámica del mercado de drogas y de los insumos para su producción.

La tabla 2 muestra los precios promedios en Colombia de la hoja de coca y tres productos derivados de esta, en el periodo de 2009 hasta 2013. En esta se observa que la hoja de coca tiene un valor muy bajo en comparación a sus productos derivados, lo cual permite deducir que para que un campesino obtenga un ingreso mensual mínimo que le permita cubrir los gastos de sostenimiento de una familia promedio (por ejemplo, de 4 o 5 personas), y compensar los esfuerzos en su trabajo, o bien debe realizar él mismo el primer proceso de extracción del alcaloide para obtener la pasta base, o bien dedicarse a cultivar muchas extensiones de coca si no tiene otra fuente de ingresos alternativa.

Tabla 2. Precios promedio de la hoja de coca y sus derivados, 2008-2013

Producto	2009		2010		2011		2012		2013	
	US\$ /kg	COL\$ /kg	US\$ /kg	COL\$ x 1000 /kg	US\$ /kg	COL\$ x 1000 /kg	US\$ /kg	COL\$ x 1000 /kg	US\$ /kg	COL\$ x 1000 /kg
Clorhidrato de Cocaína ¹	2.147	4.587	2.439	4.623	2.468	4.556	2.473	4.447	2.521	4.711
Base de cocaína ²	1.249	2.674	1.475	2.795	1.407	2.596	1.499	2.696	1.432	2.616
Pasta básica de cocaína ²	956	2.048	1.015	1.923	1.003	1.852	1.025	1.844	1.011	1.889
Hoja de cocaína ²	1,3	2,8	1,3	2,5	1,3	2,4	1,4	2,6	1,1	2

Nota: Los valores corresponden a dólares estadounidenses y pesos colombianos por kilogramo de producto. Valores de pesos colombianos (COL\$) x 1000

¹ Los precios capturados en principales ciudades. No se conocen los niveles de pureza.

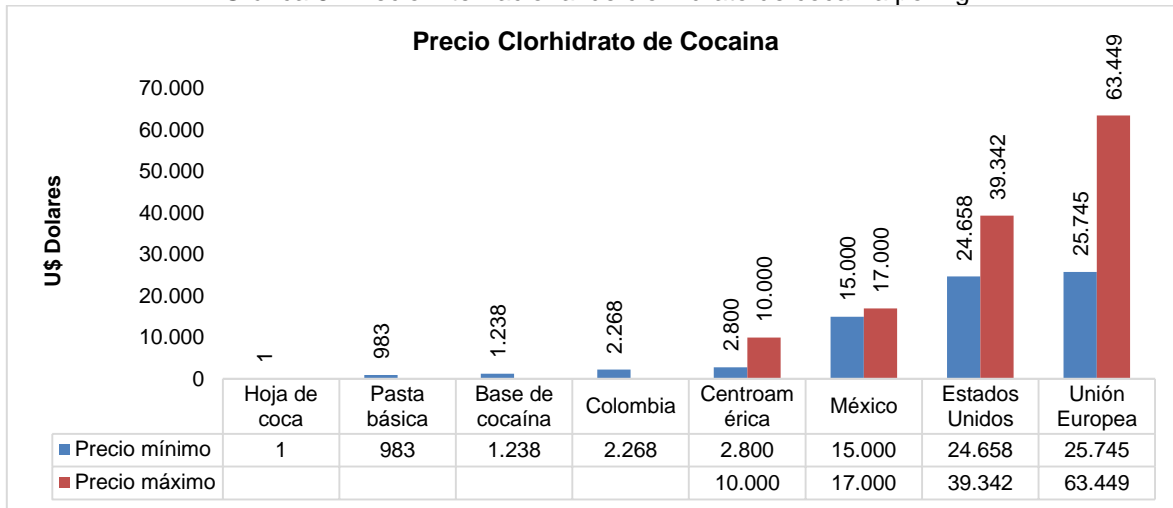
² Los precios corresponden a los registros capturados en los sitios de producción.

³ La TRM promedio para el 2013 es COL\$ 1.869

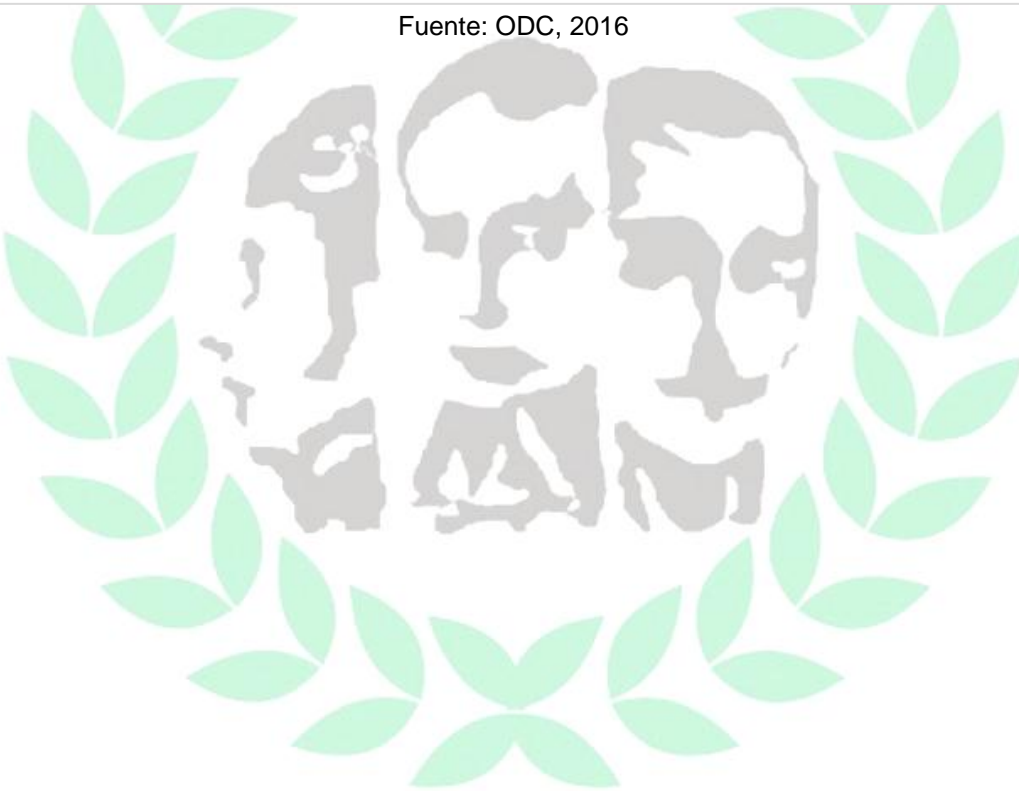
Fuente: ODC, 2016

En la gráfica 8 se pueden observar los precios de la pasta y base de cocaína y la hoja de coca en tres contextos diferentes, observándose que en Colombia están los precios más bajos de estos productos.

Gráfica 8. Precio internacional de clorhidrato de cocaína por kg



Fuente: ODC, 2016



2. MARCO TEÓRICO

2.1 ANÁLISIS DE SERIES TEMPORALES

Una serie temporal (cronológica o histórica) puede definirse como un conjunto de observaciones de una variable en distintos instantes del tiempo, en donde éstas observaciones han sido obtenidas por lo general a intervalos de tiempo regulares (Hernández, 2009).

De acuerdo a Dettling (2015), las observaciones que se han recogido sobre intervalos fijos de muestreo forman una serie cronológica. Siguiendo un enfoque estadístico, se considera que tales series son realizaciones de variables aleatorias. Una secuencia de variables aleatorias definida en intervalos fijos de muestreo, a veces se denomina *proceso estocástico de tiempo discreto*, aunque los nombres más populares son modelo de serie de tiempo o proceso de serie temporal. Este autor también resalta la importancia de diferenciar entre una serie de tiempo, es decir, los valores observados, y un proceso, es decir una construcción probabilística.

Dettling (ibid.) define formalmente un proceso de serie de tiempo como un conjunto de variables aleatorias $\{X_t, t \in T\}$, donde T es el conjunto de veces en los cuales el proceso fue, será o puede ser observado. Se asume que cada variable aleatoria X_t está distribuida de acuerdo a alguna función de distribución F_t . En esta definición se considera exclusivamente procesos de series de tiempo con intervalos de tiempo equidistantes. La tabla 3 presenta un ejemplo de una serie temporal.

Tabla 3. Serie de tiempo de número de reservas de pasajeros internacionales en la aerolínea PanAm en los Estados Unidos

Año	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1949	112	118	132	129	121	135	148	148	136	119	104	118
1950	115	126	141	135	125	149	170	170	158	133	114	140
1951	145	150	178	163	172	178	199	199	184	162	146	166
1952	171	180	193	181	183	218	230	242	209	191	172	194
1953	196	196	236	235	229	243	264	272	237	211	180	201
1954	204	188	235	227	234	264	302	293	259	229	203	229
1955	242	233	267	269	270	315	364	347	312	274	237	278
1956	284	277	317	313	318	374	413	405	355	306	271	306
1957	315	301	356	348	355	422	465	467	404	347	305	336
1958	340	318	362	348	363	435	491	505	404	359	310	337
1959	360	342	406	396	420	472	548	559	463	407	362	405
1960	417	391	419	461	472	535	622	606	508	461	390	432

*Cantidades en miles de reservas

Fuente: Dettling, M. *Applied Time Series Analysis*. Institute for Data Analysis and Process Design. Zurich University of Applied Sciences. 2015. p.(2)

El acercamiento sistemático mediante el cual se busca contestar las preguntas matemáticas y estadísticas planteadas por estas correlaciones del tiempo se conoce comúnmente como análisis de la serie de tiempo (Shumway y Stoffer, 2011). De acuerdo a estos autores, el análisis de datos experimentales que se han observado en diferentes momentos en el tiempo conduce a problemas nuevos y únicos en el modelamiento e inferencia estadísticos. El muestreo de puntos adyacentes en el tiempo introduce una correlación que puede restringir bastante la aplicabilidad de muchos métodos estadísticos convencionales, los cuales son tradicionalmente dependientes del supuesto de que estas observaciones adyacentes son independientes e idénticamente distribuidas.

Dettling (ibid.) contrasta el análisis de datos básico, donde es clave el asumir que los datos están distribuidos de forma idéntica e independiente, con el análisis de series de tiempo, donde existe una fuerte correlación. El propósito del análisis de series de tiempo es visualizar y comprender estas dependencias en datos pasados y explotarlos para predecir valores futuros.

La aplicación de algunas técnicas descriptivas sencillas a menudo permiten la comprensión de los datos, mientras que un análisis de series de tiempo completo implica el modelado del mecanismo estocástico que se supone es el generador de las series temporales observadas. Una vez que se encuentra un buen modelo y este se ajusta a los datos, el analista puede utilizar ese modelo para predecir valores futuros y producir intervalos de predicción, o también puede generar simulaciones que sirven para guiar las decisiones de planificación. Por otra parte, los modelos ajustados se utilizan como base para las pruebas estadísticas, porque sirven para determinar si, por ejemplo, las fluctuaciones en las ventas mensuales proporcionan evidencia de algún cambio subyacente, o si todavía están dentro del rango de la variación aleatoria usual (Dettling, ibid.).

Al referirse a las series temporales del campo de la economía, Hernández (ibid.) señala que, dado que los datos de una serie temporal son el resultado de una evolución histórica de un fenómeno económico concreto, con el análisis de series de tiempo se busca captar ciertas características especiales de dicho fenómeno (pautas de regularidad) con una doble finalidad práctica: describir los aspectos básicos del fenómeno en estudio y realizar predicciones de la trayectoria futura de la variable, que sean lo más acertadas posibles.

El análisis de series temporales se aplica a diversos campos de aplicación práctica en donde se presentan fenómenos importantes en el tiempo y que son muy comunes. Por ejemplo, en economía y negocios se observan semanalmente las tasas de interés, diariamente se observan cierres de precios, índices de precios mensualmente, cifras de ventas anuales. En meteorología se registran diariamente altas y bajas temperaturas, precipitaciones e índices de sequía anuales, y cada hora la velocidad del viento. En la agricultura se registran cifras anuales de producción agrícola y ganadera, erosión de los suelos y exportación de ventas. En las ciencias biológicas se observa y registra la actividad eléctrica del

corazón en intervalos de milisegundos. En ecología se registra la abundancia de una especie animal (Cryer, J.; Chan, K.; 2009).

2.1.1 Tipos de análisis de series de tiempo. De acuerdo a Hernández (ibid.), se pueden establecer históricamente tres tipos de análisis de series de tiempo: *análisis clásico de series temporales*, *análisis causal*, y *análisis moderno o enfoque Box-Jenkins* (Hernández, ibid.).

- **Análisis clásico de series temporales.** En este, la serie temporal de una variable se subdivide en tendencia, ciclo, movimiento estacional y movimiento irregular, buscando aislar cada uno de estos componentes para luego extrapolarlos hacia el futuro y obtener el pronóstico de la variable.
- **Análisis causal o enfoque estructural.** En este enfoque la evolución de la variable en el tiempo se explica a través de una ecuación que relaciona la variable estudiada con otras variables que inciden sobre ella, de forma tal que, conociendo los valores futuros de las variables causa, se determinan los valores futuros de la variable explicada.
- **Análisis moderno de series temporales o enfoque Box-Jenkins.** En este enfoque se explica la evolución de la variable temporal en función del comportamiento pasado de la propia variable, para lo cual se establece primero la estructura de dependencia entre valores sucesivos de la serie (modelo ARIMA) y luego es posible predecir la evolución futura de la variable.

2.1.2 Procesos estocásticos. Según Mascareñas (2013), se denomina proceso estocástico a toda variable que evoluciona a lo largo del tiempo de forma total o parcialmente aleatoria. Un ejemplo es la temperatura de una ciudad, en donde por lo general esta sube o aumenta en el día y disminuye en la noche, y también aumenta durante el verano y disminuye durante el invierno. La variación de temperatura es parcialmente determinística y parcialmente aleatoria.

En otra definición de proceso estocástico, Mauricio (2007) indica que es una secuencia de variables aleatorias, ordenadas y equidistantes cronológicamente, referidas a una o a varias características de una unidad observable en diferentes momentos (se llama proceso univariante o escalar cuando la secuencia se refiere a una característica y multivariante o vectorial cuando son varias).

Gujarati (2003) indica que el término *aleatorio* es sinónimo de *estocástico*, al explicar que una variable aleatoria o estocástica es aquella que puede tomar cualquier conjunto de valores, positivos o negativos, con una probabilidad dada. Este autor define un proceso estocástico o aleatorio como una colección de variables aleatorias ordenadas en el tiempo, observando que, si Y denota una variable aleatoria y es continua, se denota como $Y(t)$, pero si es discreta se expresa como Y_t (Gujarati, 2010).

Hernández (ibid.) define un proceso estocástico como una familia de variables aleatorias referidas a distintos instantes del tiempo, mientras que una serie de tiempo es una muestra de tamaño I del proceso estocástico con una observación para cada periodo. En otras palabras, un proceso estocástico (población) genera la serie temporal (muestra), por lo cual se asume que a partir de esta última se puede determinar cuál es el proceso que ha generado la serie muestral.

Se puede representar entonces

Y_1, Y_2, \dots, Y_T
Proceso estocástico

Y_1, Y_2, \dots, Y_T
Serie temporal

Conocer el proceso que ha generado la serie temporal supone entonces que es necesario reconocer la función de distribución conjunta del proceso, basándose en la información que proporciona la serie; se trata entonces de conocer

$$P(Y_1, Y_2, \dots, Y_T)$$

2.2 MODELOS ESTADÍSTICOS DE SERIES DE TIEMPO

2.2.1 Componentes inobservados de las series de tiempo. Debido a que los componentes no pueden observarse directamente a través de los datos de la serie temporal (por lo cual a estos métodos también se les conoce como enfoque de componentes inobservados), es necesario aproximarse a ellos por medio de procedimientos matemático-estadísticos que no siempre garantizan una eficaz descomposición de la serie temporal.

2.2.2 Modelo clásico de series de tiempo. En el análisis clásico, las variaciones en el tiempo que presentan las series temporales se explican por la integración de cuatro componentes básicos: *tendencia (T)*, *ciclo (C)*, *estación (E)*, y *residuo (R)* (Hernández, ibid.).

- **Tendencia:** es la trayectoria a largo plazo que sigue la serie. Por ejemplo, la gráfica 9 muestra la serie temporal del consumo mensual de gasolina en España entre enero de 1966 y agosto de 1977. En esta se observa que la serie presenta un comportamiento general de incremento constante, de tipo lineal, a medida que transcurre el tiempo. Como la gráfica de la serie representa los datos reales, contiene todos los componentes, por lo cual no se pueden identificar a simple vista los componentes. Sin embargo, en la gráfica 10, al hacer la descomposición gráfica de la serie, la tendencia (trend) es fácilmente observable, y se destaca su incremento positivo de tipo lineal.
- **Ciclo:** Esta componente refleja comportamientos recurrentes (que no son necesariamente periódicos) con un periodo superior a un año. Muestran, habitualmente, cómo se suceden las etapas de bonanza económica con las de crisis, o al menos, desaceleración. En series económicas lo conforman las

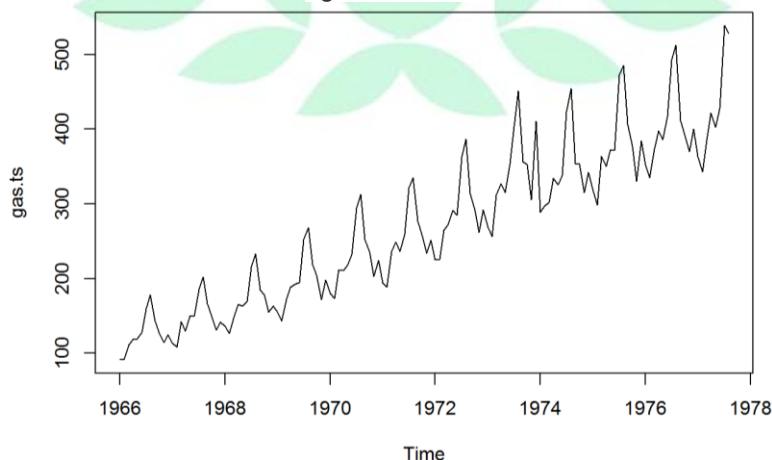
fluctuaciones de la serie que se reproducen a mediano plazo (5 a 8 años), además de que en las series económicas, muchas veces los ciclos no son fácilmente reconocibles, y por tanto no se consiguen separar de la tendencia, al menos para series no demasiado largas.

- Variaciones estacionales: son las variaciones que se producen en periodos con duración de un año o menos y que son generadas por causas climáticas o usos sociales. En el caso de series económicas, muchas series presentan oscilaciones regulares en el mismo mes de cada año, y con unas pautas que se presentan, sin repetirse exactamente, todos los años, las cuales se deben básicamente a causas climatológicas, vacacionales o fiscales.

Sin embargo, la estacionalidad no se presenta sólo cuando el periodo amplio es el año, ya que pueden ocurrir estacionalidades mensuales o semanales en series diarias, o estacionalidades diarias en series horarias (ej. las series de cotizaciones bursátiles). Es importante que las estacionalidades tengan un periodo no superior al anual, para que no se confundan con las componentes cíclicas. En la gráfica 10 se observa claramente el componente estacional (seasonal) de la serie del consumo mensual de gasolina en España entre enero de 1966 y agosto de 1977. Se observa que la serie presenta un comportamiento que se repite aproximadamente en periodos de un año, y que tiene sus puntos más altos hacia la mitad (junio) de cada año.

- Variaciones residuales o residuo: son movimientos erráticos de la serie, que no muestran un carácter periódico reconocible, y que se atribuyen a factores no determinados o al azar. Esta componente irregular recoge, en consecuencia, la incapacidad del modelo de la serie para explicar a la perfección el comportamiento de la serie temporal. En la parte inferior de la gráfica 10 se observa la componente residual o aleatoria (random) de la serie de tiempo ya mencionada.

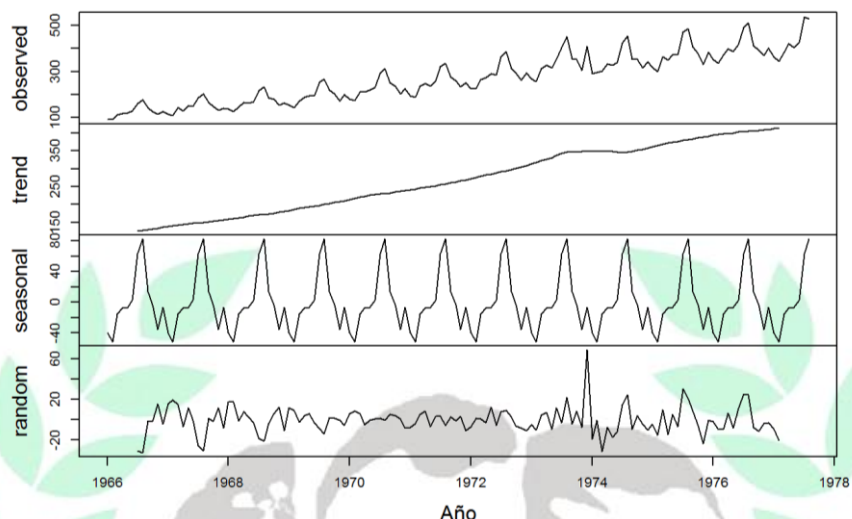
Gráfica 9. Serie de tiempo del consumo mensual de gasolina en España entre enero de 1966 y agosto de 1977



Fuente: Berrendero, J. (2017). *Análisis básico de series temporales con R*. R Pubs brought to you by RStudio. En: <https://rpubs.com/joser/SeriesTemporalesBasicas>

En la gráfica 10 se muestra la descomposición gráfica obtenida con R de la serie de tiempo anterior.

Gráfica 10. Serie original del consumo mensual de gasolina en España (enero de 1966 hasta agosto de 1977) y su descomposición en componentes aditivos



Fuente: Berrendero, J. (2017). *Análisis básico de series temporales con R*. R Pubs brought to you by RStudio. En: <https://rpubs.com/joser/SeriesTemporalesBasicas>.

De acuerdo a Hernández (ibid.), los componentes inobservados se integran a través de dos hipótesis básicas: la *hipótesis aditiva*, que asume que cada valor de la serie de tiempo en un momento t es el resultado de sumar los valores de cada componente individual para ese momento (aunque no siempre aparecen todos los componentes), y la hipótesis *multiplicativa*, que asume que cada valor de la serie resulta del producto de los componentes individuales. Así:

$$Y_t = T_t + C_t + E_t + R_t \quad (\text{hipótesis aditiva})$$

$$Y_t = T_t \times C_t \times E_t \times R_t \quad (\text{hipótesis multiplicativa})$$

Otra tercera hipótesis se llama *modelo multiplicativo linealizado*, que consiste en que un modelo multiplicativo de trabaja con los datos de la serie temporal en logaritmos y no con los datos originales de la variable. Así, el modelo está representado por:

$$\ln Y_t = \ln T_t + \ln C_t + \ln E_t + \ln R_t$$

En este caso, después de realizar el análisis y agregar los componentes de acuerdo al esquema aditivo, se calcula el antilogaritmo para obtener el valor de la variable de interés.

2.2.3 Modelos ARIMA (enfoque Box-Jenkins). La designación de modelo ARIMA proviene de “AutoRegresive Integrated Moving Average”, y se refiere a un tipo de modelos autoprotectivos que permiten predecir los valores de una variable en función de sus valores pasados sin necesidad de otra información de variables

auxiliares o relacionadas. El nombre genérico ARIMA de estos modelos hace referencia a sus tres componentes principales: Autoregresivo (AR), Integrado (I), de Medias Móviles (MA). En estos modelos, cada observación de una variable en un momento dado es modelada en función de valores anteriores en el tiempo para la misma variable (Pérez, C.; 2006).

El este enfoque, el proceso de modelizar una serie temporal consiste en derivar un modelo ARIMA que se ajuste al conjunto de datos dado, para lo cual, se requiere realizar un proceso de análisis de las características esenciales de las series de tiempo como la tendencia, estacionalidad, variaciones cíclicas, funciones de autocorrelación, residuales.

El modelo ARIMA se expresa en una ecuación explícita que permite describir un valor como una función lineal de datos anteriores y errores debidos al azar, permitiendo además poder incluir un componente cíclico o estacional de la serie de tiempo respectiva. El objetivo central consiste en obtener un modelo adecuado, pero parsimonioso, esto es, el modelo ARIMA debe contener todos los elementos necesarios, pero a la vez los mínimos que se requieren para describir el fenómeno. Box y Jenkins (citados por Pérez, C.; 2006) recomiendan que se deben tener como mínimo unos cincuenta datos de la serie temporal.

2.2.4 Evaluación de la capacidad predictiva en modelos ARIMA. De acuerdo a De Arce y Mahía (s.f.), los criterios para seleccionar los modelos con mejor capacidad predictiva se basan en la premisa de que una predicción es mejor que otra cuando comete menor error; estos criterios son el *error cuadrático medio (ECM)*, el *error absoluto medio (EAM)*, y el *error absoluto porcentual medio (EAPM)*.

Por su parte, Mauricio (2007) define los criterios de la siguiente forma:

- La raíz del error cuadrático medio (RMSE, del inglés Root Mean Squared Error) es la desviación típica muestral de los errores de predicción.
- El error absoluto medio (MAE, del inglés Mean Absolute Error) es la media de los valores absolutos de los errores de predicción.
- El error porcentual absoluto medio (MAPE, del inglés Mean Absolute Percentage Error) es la media de los valores absolutos de los errores de predicción porcentuales.

3. MARCO METODOLÓGICO

Para el desarrollo del proyecto propuesto, se considera necesario en primer lugar hacer una revisión exhaustiva de las bases de datos existentes con la información de interés para los análisis; estas bases de datos contienen los registros históricos de los resultados operacionales de las cuatro Fuerzas Navales de la Armada Nacional, y que a su vez han sido las únicas variables que se han utilizado hasta ahora para realizar el proceso de proyección de metas operacionales para cada siguiente año (ej. incautación marihuana, incautación insumos sólidos, incautación insumos líquidos, erradicación cultivos ilícitos).

Se hace necesario también, hacer una clasificación de las variables del fenómeno delictivo del narcotráfico reconocidos normativamente, asociando a ellos resultados operacionales por años, con lo cual se pretenden conocer tendencias en la presentación del fenómeno y de los enfoques en los esfuerzos hechos por las Fuerzas Navales. La clasificación del fenómeno delictivo del narcotráfico debe corresponder a la establecida en el Código Penal colombiano, ya que, tanto la Armada Nacional como las demás instituciones del Estado, se rigen por esta tipología.

De esta forma, se han considerado las siguientes fases para el desarrollo del proyecto:

Fase 1.

- Analizar y consultar permanente la base de datos del Sistema de Información Geográfico Operacional SIGO que contienen la información histórica de los resultados operacionales de la Armada Nacional (Indicadores de resultados), variables asociadas al fenómeno delictivo del narcotráfico.
- Encontrar datos atípicos y vacíos de datos.

Fase 2.

- Clasificar y organizar las variables asociadas al fenómeno delictivo del narcotráfico de forma independiente, (Indicadores de resultados) correspondientes a cada vigencia.

Fase 3.

- Realizar análisis de estadística descriptiva para las variables del narcotráfico de acuerdo a cada año de registros.
- Identificar y condensar los cambios o variaciones en el tiempo, inicialmente por años, de la dinámica de las variables del narcotráfico (resultados operacionales registrados históricamente).

Fase 4.

- Realizar análisis gráfico (geo-referenciado) sobre la cartografía colombiana y sobreponer las variables del fenómeno delictivo del narcotráfico para verificar como confluyen estos en las diferentes zonas del país.

Fase 5.

- Elaborar modelos estadísticos de líneas de series de tiempo, a partir de los registros por años.
- Realizar pruebas de análisis de tipo predictivo con los modelos estadísticos definidos.

3.1 RECURSOS

Se cuenta con una base de datos bastante robusta, donde se tiene información de resultados operacionales desde el año 2003, con nivel de detalle (año-mes--días); se tiene disponible la plataforma informática del Sistema de Información Geográfico Operacional (SIGO), el Sistema de Información de Geoposicionamiento (CAPRICORNIO). Se requiere además, el software estadístico R-Studio versión 1.0.136, el software Latex versión en línea, dos (2) equipos de cómputo, una (1) impresora, un (1) asesor metodológico de la Fundación Universitaria los Libertadores, dos (2) asesores en estadística de la Fundación Universitaria los Libertadores, el equipo de investigadores que está definido por nuestro equipo de trabajo (dos (2) investigadores).

Los análisis se realizarán con el aplicativo informático R-Studio, versión 1.0.136

3.2 IMPACTOS

Los impactos esperados en el desarrollo de este proyecto de investigación se resumen como:

Impacto esperado	Supuesto de desarrollo	Beneficiarios
Replicar el modelo en las unidades operativas menores.	Modelo en las unidades operativas menores tales como las cinco Brigadas de Infantería de Marina.	Brigadas de Infantería de Marina, Flotillas de Superficie, Guardacostas y Batallones de Infantería de Marina.
Metas operacionales para cada una de las variables, en las cinco Brigadas de Infantería de Marina de la Armada Nacional de Colombia.	Metas para cada variable (20) por cada Brigada de Infantería de Marina (05)	Brigadas de Infantería de Marina, Flotillas de Superficie, Guardacostas y Batallones de Infantería de Marina.

4. ANÁLISIS Y RESULTADOS

4.1 SERIE INCAUTACIONES DE COCAÍNA 2012-2016

La serie de tiempo obtenida a través de la Dirección de Operaciones Navales está conformada por los registros de incautaciones de clorhidrato de cocaína en miles de kilogramos (kg) realizadas por la Armada Nacional en la jurisdicción y áreas de responsabilidad fluvial que tiene a cargo en todo el territorio nacional. La tabla 4 muestra la serie de tiempo mensual que comprende el periodo que inicia en enero de 2012 hasta el mes de diciembre de 2016, incluyendo en total 60 datos.

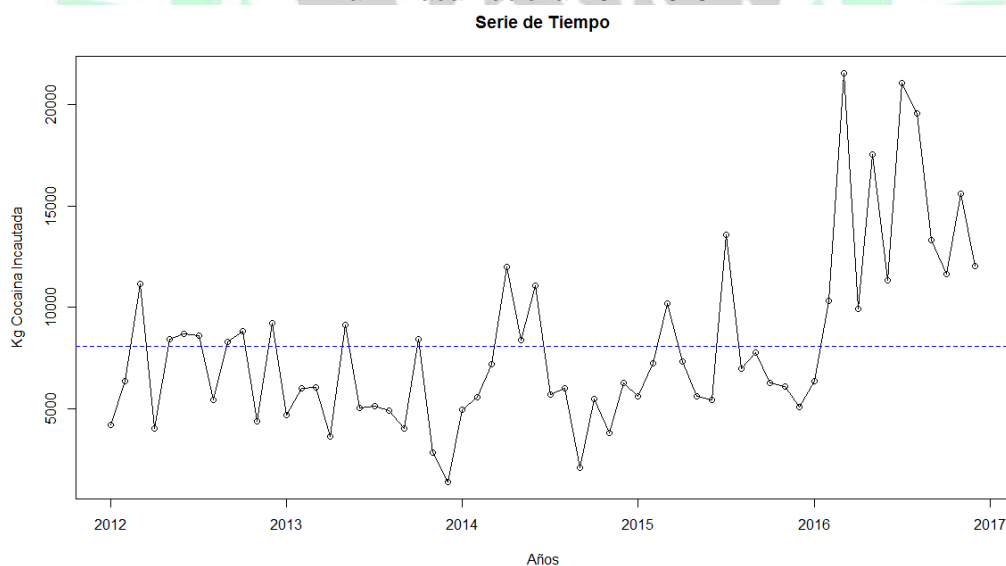
Tabla 4. Serie de tiempo de incautaciones de cocaína en kg realizadas por la Armada Nacional

Años	ene	feb	mar	abr	may	jun	jul	ago	sep	oct	nov	dic
2012	4.201	6.355	11.177	4.039	8.426	8.719	8.618	5.449	8.323	8.821	4.403	9.224
2013	4.696	6.031	6.071	3.654	9.139	5.064	5.148	4.931	4.054	8.439	2.863	397
2014	4.984	5.574	7.215	11.997	8.405	11.090	5.714	6.007	2.117	5.490	3.835	6.282
2015	5.631	7.267	10.208	7.341	5.629	5.438	13.571	7.001	7.777	6.282	6.110	5.096
2016	6.377	10.327	21.551	9.916	17.520	11.328	21.060	19.571	13.337	11.643	15.609	12.043

Fuente: Dirección de Operaciones Navales. Bogotá. 2017

El comportamiento de la serie de incautaciones de cocaína se muestra en la gráfica 11. Con esta ayuda visual se puede apreciar que la serie presenta un cambio en el nivel a partir del año 2016, pero no se ve claramente si hay o no tendencia en la serie.

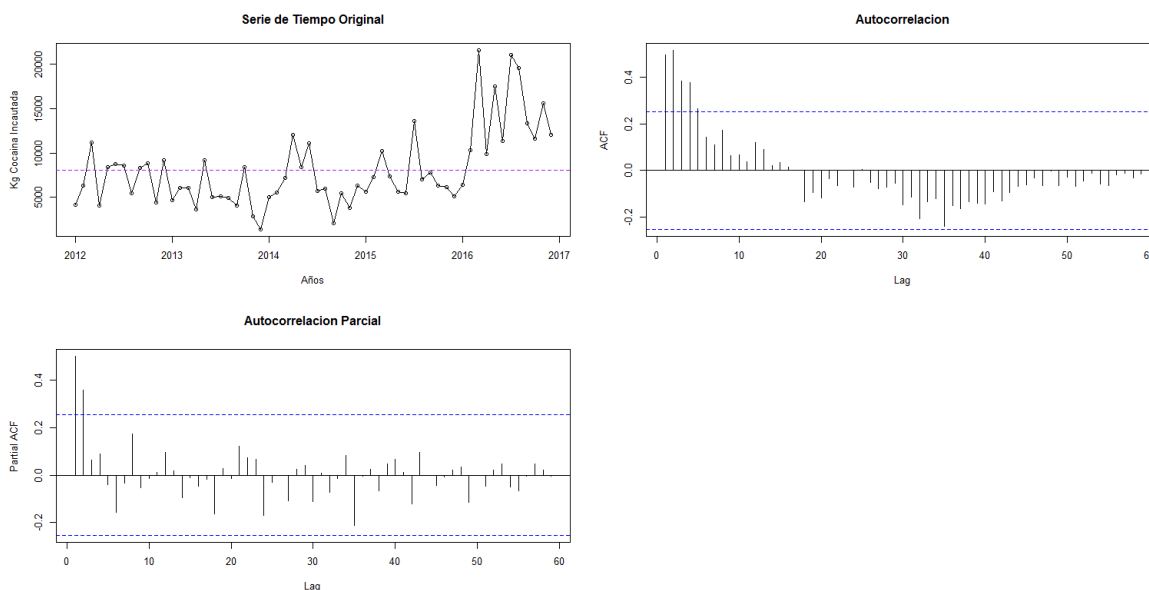
Gráfica 11. Serie de tiempo de incautaciones mensuales de kilogramos de cocaína realizadas por la Armada Nacional 2012-2016



Fuente: elaboración propia (2017)

La gráfica múltiple 12 muestra la serie original junto a los correlogramas ACF y PACF. En el gráfico de auto correlación (ACF) se observa la existencia de tendencia en la serie.

Gráfica 12. Serie de tiempo de incautaciones de cocaína (kg) y correlogramas



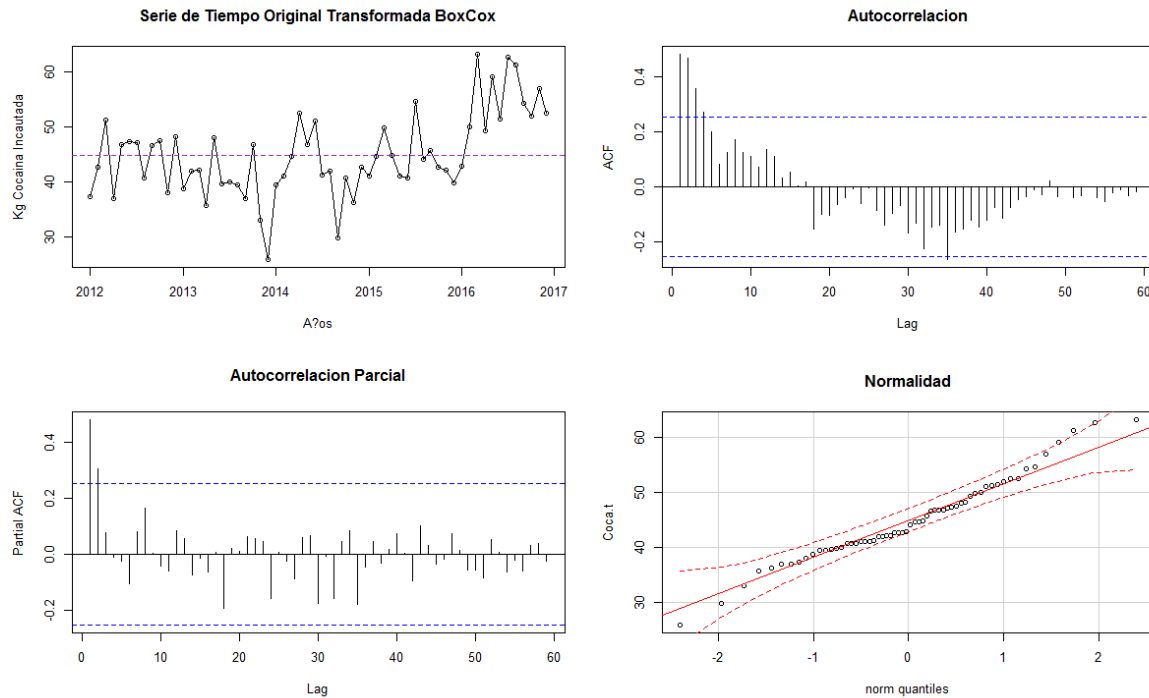
Fuente: Elaboración propia (2017)

Para comprobar si la serie es o no estacionaria, se realiza la prueba “Dickey-Fuller Test”. Realizando la prueba en el aplicativo “R Studio”, el valor obtenido del $p\text{-value}=0.4382$ es mayor que $\alpha=0.05$, lo que quiere decir que se acepta H_0 (la serie no es estacionaria) y se rechaza la hipótesis alterna H_1 (la serie es estacionaria). Por tanto, la serie es no estacionaria.

La serie de tiempo mantiene un problema de falta de normalidad; por esto se realiza la transformación de Box-Cox en el aplicativo “R Studio”.

La gráfica múltiple 13 muestra la serie de tiempo transformada con BoxCox junto a la gráfica de ACF y PACF, en las cuales se observa que la serie aun presenta tendencia, pero mejora mucho la normalidad.

Gráfica 13. Serie de tiempo transformada BoxCox de cocaína (kg)

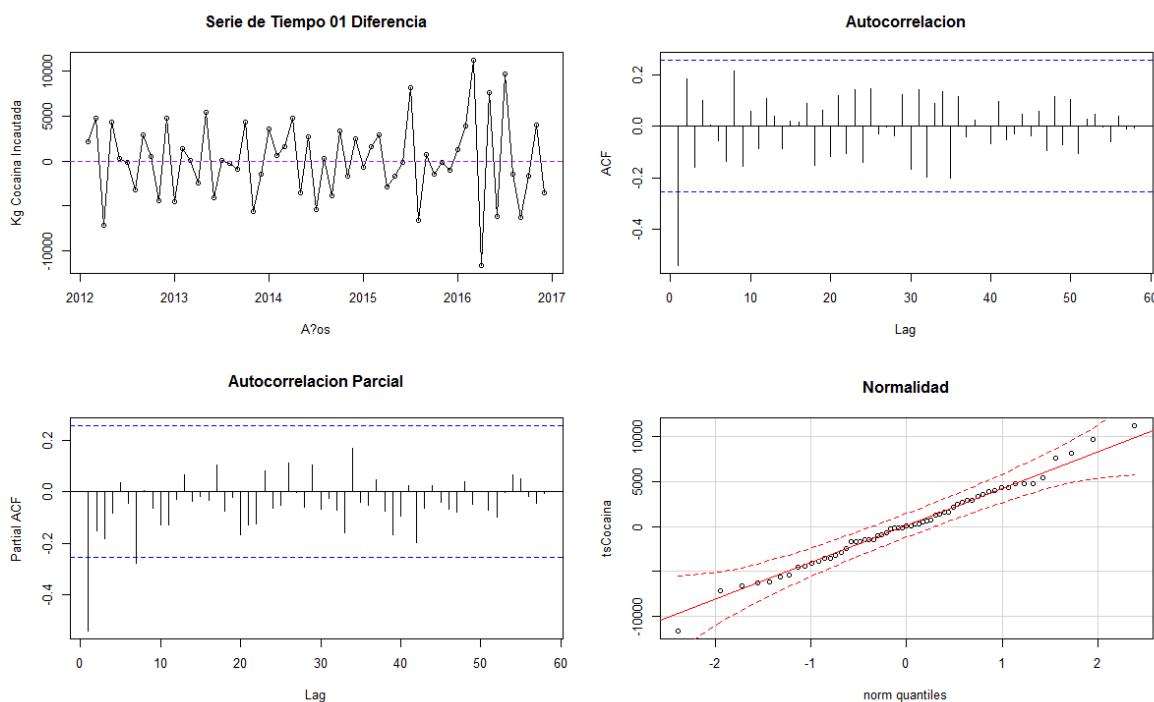


Fuente: Elaboración propia (2017)

Para volver la serie estacionaria se realiza una diferencia en la serie de tiempo ya transformada. Se realiza entonces la prueba de “Dickey-Fuller Test” para la serie con 01 diferencia. El valor obtenido del p-value es menor que $\alpha = 0.05$, lo que quiere decir que se rechaza H_0 y se acepta la hipótesis alterna H_1 , que indica que la serie ya es estacionaria.

La gráfica múltiple 14 muestra la serie de tiempo transformada con BoxCox y una (01) diferencia junto a la gráfica de ACF y PACF, en las cuales se observa que la serie ya no presenta tendencia y también mantiene una estructura latente de normalidad.

Gráfica 14. Serie de tiempo de incautaciones de cocaína transformada con BoxCox y una (01) diferencia



Una vez comprobada la estacionalidad de la serie, se examinan los gráficos correlogramas ACF y PACF; estos son herramientas comúnmente usadas para la identificación de la relación mutua existente entre los valores de la serie de tiempo en diferentes periodos y describe lo que tiende a suceder a un valor si se da un cambio en el otro. En el gráfico de autocorrelación simple (ACF) se evidencia que los rezagos (Lag) número uno (1) se sale de las bandas de confianza, al igual que en el gráfico de autocorrelación parcial (PACF) los rezagos 1 y 7 se salen de las bandas también.

A partir de este análisis se proponen los siguientes modelos a probar:

ARIMA(4.1.0), ARIMA(0.1.4), ARIMA(1.1.4), ARIMA(7.1.1), ARIMAX(2.1.4), ARIMAX(4.1.1), ARIMAX(1.1.7), ARIMAX(3.1.1).

4.2 PRUEBAS DE LOS MODELOS ARIMA Y ARIMAX INICIALES

Se procedió a realizar las pruebas de capacidad predictiva de los modelos ARIMA en R-studio; la tabla 5 reúne los resultados de las pruebas de normalidad, de correlación y del BIC para los ocho (8) modelos probados.

Tabla 5. Comparativo de las pruebas de los modelos ARIMA Y ARIMAX iniciales

Análisis Modelos	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6	Modelo 7	Modelo 8
	ARIMA (4.1.0)	ARIMA (0.1.4)	ARIMA (1.1.4)	ARIMA (7.1.1)	ARIMAX (2.1.4)	ARIMAX (4.1.1)	ARIMAX (1.1.7)	ARIMAX (3.1.1)
Criterio Información Bayesiana (BIC)	BIC 1144	BIC 1198	BIC 1159	BIC 1159	BIC 1197	BIC 1159	BIC 1198	BIC 1153
Normalidad Prueba Jarque Bera Test	12,52%	5,94%	0,31%	0,34%	0,39%	0,31%	5,94%	0,00%
Correlacion Prueba Ljung-Box Test	83,31%	1,06%	0,04%	0,27620	0,02%	0,04%	1,06%	79,72%

Fuente: Elaboración propia (2017)

4.3 PRUEBAS DE LOS MODELOS ARIMAX SIN ANÁLISIS DE INTERVENCIÓN

Se procedió a realizar también las pruebas de capacidad predictiva de los modelos ARIMAX con variables regresoras incorporadas. La tabla 6 reúne los resultados de las pruebas de normalidad, de correlación y del BIC para los ocho (8) modelos ARIMAX que incorporan variables regresoras.

Tabla 6. Comparativo de las pruebas de los modelos ARIMAX sin análisis de intervención

Análisis Modelos	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6	Modelo 7	Modelo 8
	ARIMAX (2.1.4)	ARIMAX (0.1.1)	ARIMAX (1.1.0)	ARIMAX (3.1.0)	ARIMAX (4.1.0)	ARIMAX (7.1.0)	ARIMAX (6.1.2)	ARIMAX (2.1.4)
Criterio Información Bayesiana (BIC)	BIC 1143	BIC 1131	BIC 1139	BIC 1136	BIC 1136	BIC 1141	BIC 1125	BIC 1128
Normalidad Prueba Jarque Bera Test	0,00%	0,01%	1,79%	0,03%	0,01%	0,01%	0,01%	0,04%
Correlacion Prueba Ljung-Box Test	90,28%	80,17%	76,99%	76,12%	74,99%	99,53%	95,61%	95,89%

Fuente: Elaboración propia (2017)

Para elegir el modelo más adecuado que se adapte a la serie de tiempo de las incautaciones de cocaína, los criterios que se toman en cuenta son:

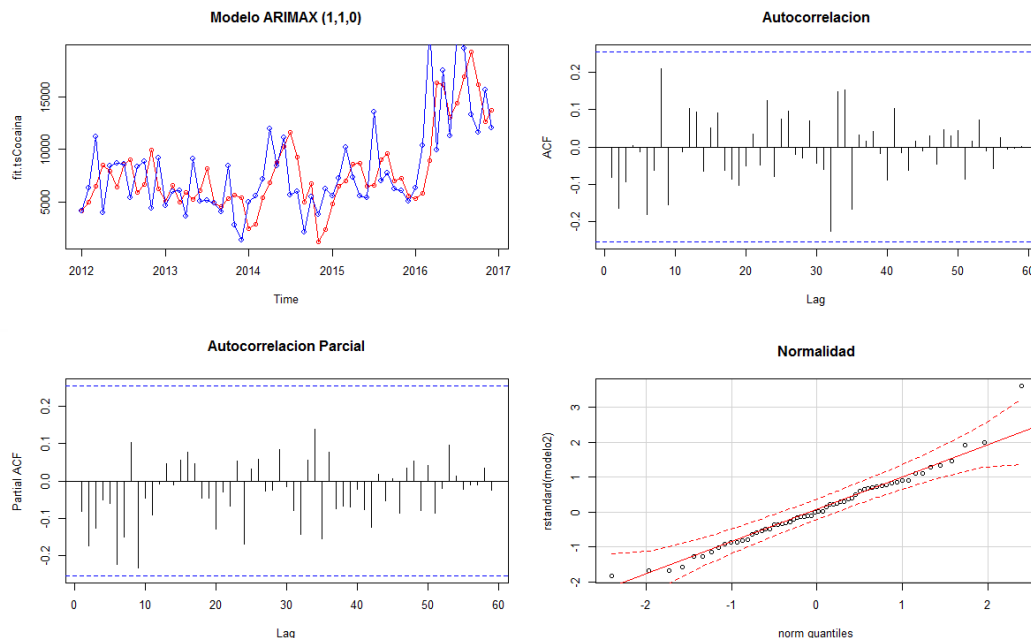
- El que tenga el menor valor del “Criterio Información Bayesiana” (BIC)
- El que la prueba “Jarque-Bera” demuestre mejor normalidad.
- El que la prueba de independencia “Ljung-Box” indique que los residuales no presenten correlación serial.

En general, de los dieciséis (16) modelos estadísticos establecidos se puede observar que los modelos que más se ajustan a los criterios anteriormente relacionados son los modelos ARIMAX. De igual forma, con base en los mismos criterios se puede inferir que los modelos que tentativamente se ajustan mejor son el ARIMAX (1.1.0) con 01 variable regresora, el ARIMAX (6.1.2) con 01 variable regresora y el ARIMAX (2.1.4) con 02 variables regresoras (ver tabla 6).

Con base en lo anterior, a continuación se procede a ilustrar las predicciones de los últimos (3) tres modelos: ARIMAX (1.1.0) con 01 variable regresora, el ARIMAX (6.1.2) con 01 variable regresora y el ARIMAX (2.1.4) con 02 variables regresoras.

4.3.1 Prueba del modelo ARIMAX (1.1.0) con 01 regresor. La gráfica múltiple 15 presenta la superposición de la serie original (en color azul) y la serie ajustada del modelo ARIMAX (1.1.0) con 01 regresor (en color rojo), los correlogramas de esta serie ajustada y el gráfico de la prueba de normalidad. El regresor introducido al modelo corresponde a la variable *Gasolina*, que fue la que resultó ser significativa en el modelo ARIMAX que se probó inicialmente con 8 variables regresoras.

Gráfica 15. Prueba del modelo ARIMAX (1.1.0) con 01 regresor



Fuente: Elaboración propia (2017)

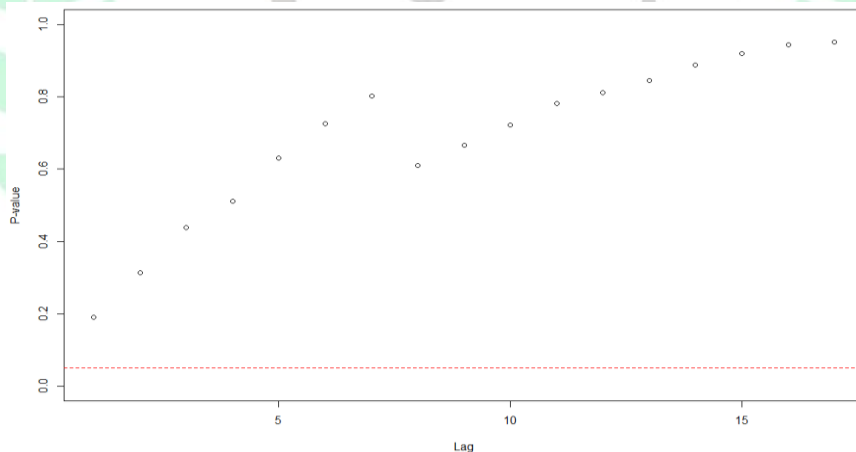
En el gráfico de autocorrelación simple (ACF) se evidencia que todos los rezagos (Lag) se encuentran dentro de la banda de confianza, al igual que en el gráfico de autocorrelación parcial (PACF). En el gráfico de normalidad se muestra que un (1) punto se sale de las bandas de confianza, evidenciando que la serie de tiempo no sigue una relación normal.

Se realiza entonces la prueba de normalidad del modelo a través de la prueba “Jarque Bera Test”, obteniéndose un p-value = 0.0179, el cual es menor que $\alpha = 0.05$, lo que quiere decir que se acepta H_0 y se rechaza la hipótesis alterna H_1 y se concluye que los residuales no siguen un comportamiento normal.

Luego se verifica la correlación de los residuales a través de la prueba “Ljung-Box”, obteniéndose un p-value = 76.99%, que es mayor que $\alpha = 0.05$, por lo cual se rechaza H_0 (hay correlación) y se acepta H_1 (no hay correlación), lo que quiere decir que los residuales son independientes uno del otro.

La gráfica 16 muestra los residuales (puntos) por encima de la banda de confianza, demostrando la homocedasticidad del modelo ARIMAX (1.1.0) con 01 regresor.

Gráfica 16. Prueba de homocedasticidad del modelo ARIMAX (1.1.0) con 01 regresor



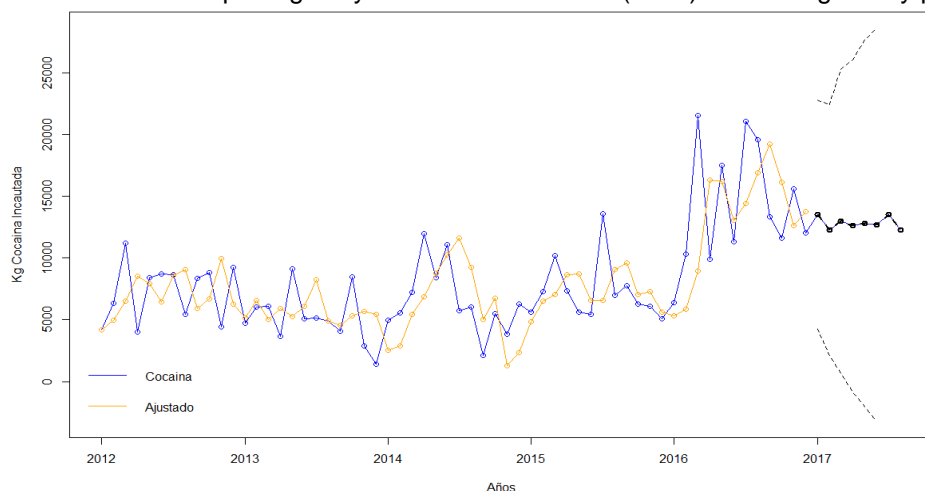
Fuente: Elaboración propia (2017)

Respecto a la prueba de homocedasticidad, se concluye que, con un nivel de significancia del 5%, hay suficiente evidencia estadística para afirmar que 95% de los residuales son homocedásticos, o sea, la varianza es constante para los primeros 20 rezagos.

La gráfica 17 muestra la superposición de la serie original (en color azul) y la serie ajustada del modelo ARIMAX (1.1.0) con 01 regresor (en color amarillo). En la

parte final de la serie original se observan los pronósticos hechos con este modelo para los primeros 8 meses del año 2017.

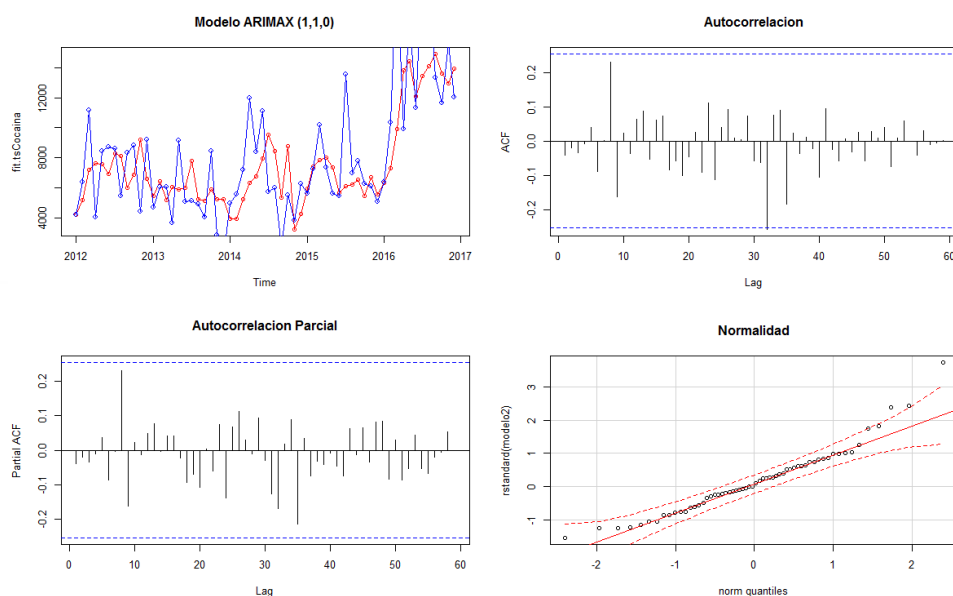
Gráfica 17. Serie de tiempo original y del modelo ARIMAX (1.1.0) con 01 regresor y pronóstico



Fuente: Elaboración propia (2017)

4.3.2 Prueba del modelo ARIMAX (6.1.2) con 01 regresor. La gráfica múltiple 18 presenta la superposición de la serie original (en color azul) y la serie ajustada del modelo ARIMAX (6.1.2) con 01 regresor (en color rojo), los correlogramas de esta serie ajustada y el gráfico de la prueba de normalidad. El regresor introducido a este modelo corresponde también a la variable *Gasolina*, que fue la que resultó ser significativa en el modelo ARIMAX que se probó inicialmente con 8 variables regresoras.

Gráfica 18. Prueba del modelo ARIMAX (6.1.2) con 01 regresor



Fuente: Elaboración propia (2017)

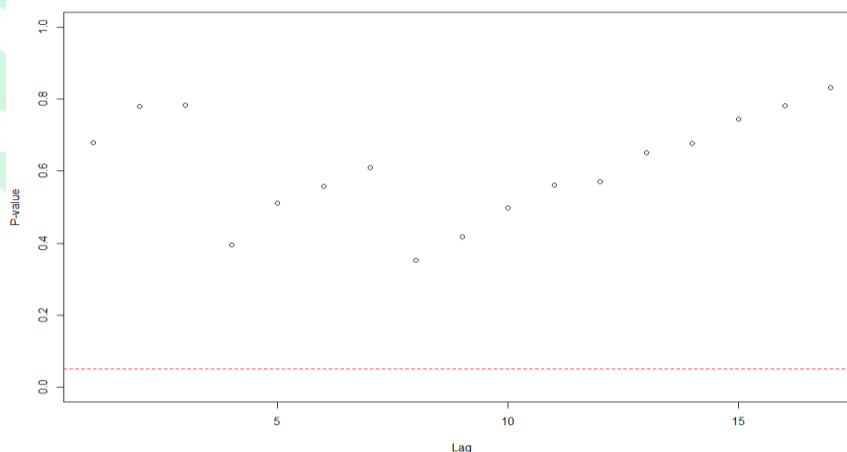
En el gráfico de autocorrelación simple (ACF) se evidencia que todos los rezagos (Lag) se encuentran dentro de la banda de confianza, al igual que en el gráfico de autocorrelación parcial (PACF). En el gráfico de normalidad se muestra que dos (3) puntos se salen de las bandas de confianza, evidenciando que la serie de tiempo no sigue una relación normal.

Se realiza entonces la prueba de normalidad del modelo a través de la prueba “Jarque Bera Test”, obteniéndose un $p\text{-value} = 0.01$, el cual es menor a $\alpha = 0.05$ lo que quiere decir que se acepta H_0 y se rechaza la hipótesis alterna H_1 y se concluye que los residuales no siguen un comportamiento normal.

Luego se verifica la correlación de los residuales a través de la prueba “Ljung-Box”, obteniéndose un $p\text{-value} = 99.53\%$, que es mayor a $\alpha = 0.05$, por lo cual se rechaza H_0 (hay correlación) y se acepta H_1 (no hay correlación), lo que quiere decir que los residuales son independientes uno del otro.

La gráfica 19 muestra que los residuales (puntos) están por encima de la banda de confianza, evidenciando la homocedasticidad del modelo ARIMAX (6.1.2) con 01 regresor.

Gráfica 19. Prueba de homocedasticidad del modelo ARIMAX (1.1.0) con 01 regresor



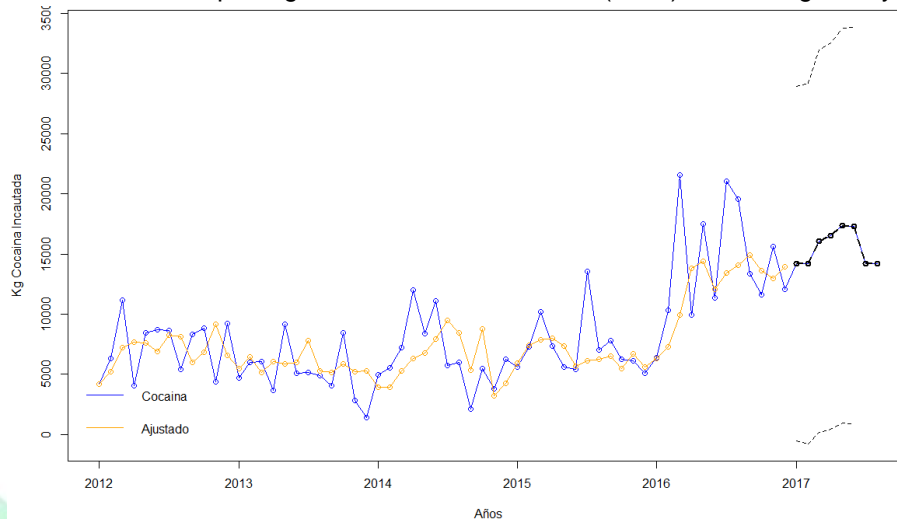
Fuente: Elaboración propia (2017)

Respecto a la prueba de homocedasticidad de este modelo, se concluye que, con un nivel de significancia del 5%, hay suficiente evidencia estadística para afirmar que el 93% de los residuales son homocedásticos, o sea, la varianza es constante para los primeros 20 rezagos.

A continuación, la gráfica 20 muestra la superposición de la serie original (en color azul) y la serie ajustada del modelo ARIMAX (6.1.2) con 01 regresor (en color

amarillo). En la parte final de la serie original se observan los pronósticos hechos con este modelo para los primeros 8 meses del año 2017.

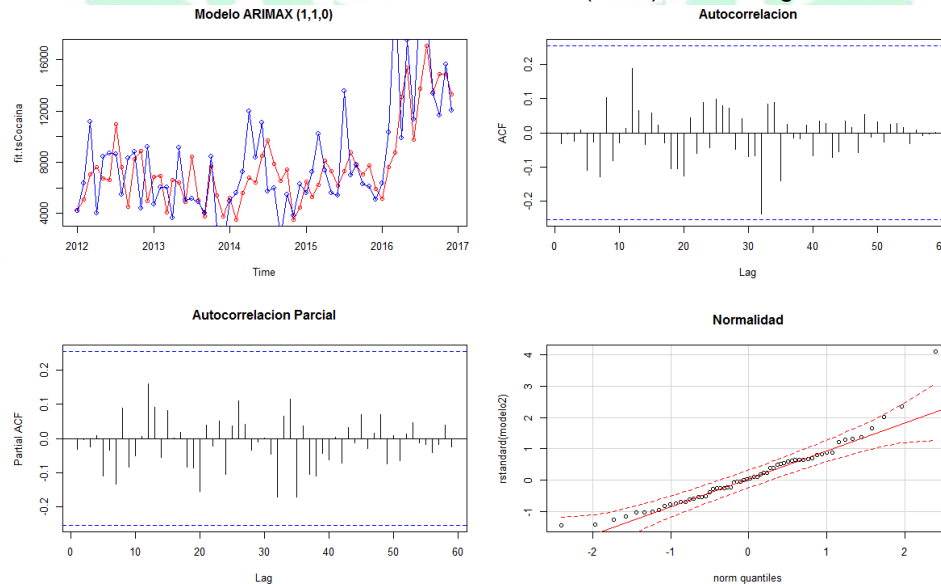
Gráfica 20. Serie de tiempo original, del modelo ARIMAX (6.1.2) con 01 regresor y pronóstico



Fuente: Elaboración propia (2017)

4.3.3 Prueba del modelo ARIMAX (2.1.4) con 02 regresores. La gráfica múltiple 21 presenta la superposición de la serie original (en color azul) y la serie ajustada del modelo ARIMAX (2.1.4) con 02 regresores (en color rojo), los correlogramas de esta serie ajustada y el gráfico de la prueba de normalidad. Los regresores introducidos al modelo corresponden a las variables *Gasolina* e *Insumos Líquidos*, que fueron los que resultaron ser más significativos en el modelo ARIMAX que se probó inicialmente con las 8 variables regresoras.

Gráfica 21. Prueba del modelo ARIMAX (2.1.4) con 02 regresores



Fuente: Elaboración propia (2017)

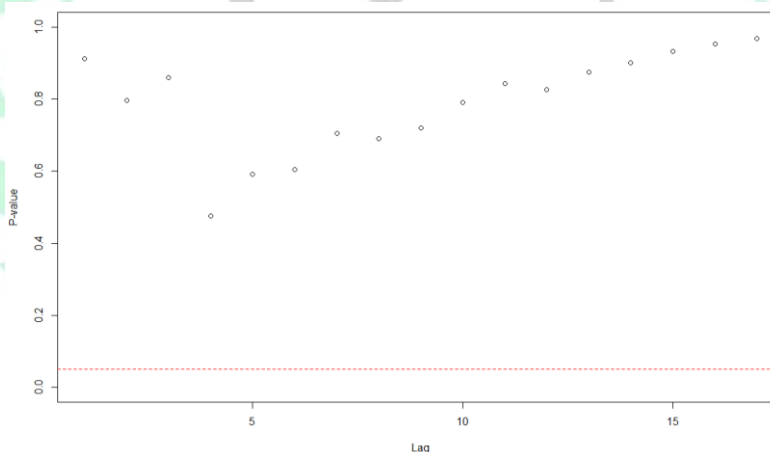
En el gráfico de autocorrelación simple (ACF) se evidencia que todos los rezagos (Lag) se encuentran dentro de la banda de confianza, al igual que en el gráfico de autocorrelación parcial (PACF). En el gráfico de normalidad se muestra que un (1) punto se sale de las bandas de confianza, evidenciando que la serie de tiempo no sigue una relación normal.

Se realiza entonces la prueba de normalidad del modelo a través de la prueba “Jarque Bera Test”, obteniéndose un p-value = 0.04, el cual es menor a $\alpha = 0.05$, lo que quiere decir que se acepta H_0 y se rechaza la hipótesis alterna H_1 y se concluye que los residuales no siguen un comportamiento normal.

Luego se verifica la correlación de los residuales a través de la prueba “Ljung-Box”, obteniéndose un p-value = 95.89%, que es mayor a $\alpha = 0.05$, por lo cual se rechaza H_0 (hay correlación) y se acepta H_1 (no hay correlación), lo que quiere decir que los residuales son independientes uno del otro.

La gráfica 22 muestra que los residuales (puntos) por encima de la banda de confianza demostrando la homocedasticidad del modelo ARIMAX (2.1.4) con 02 regresores.

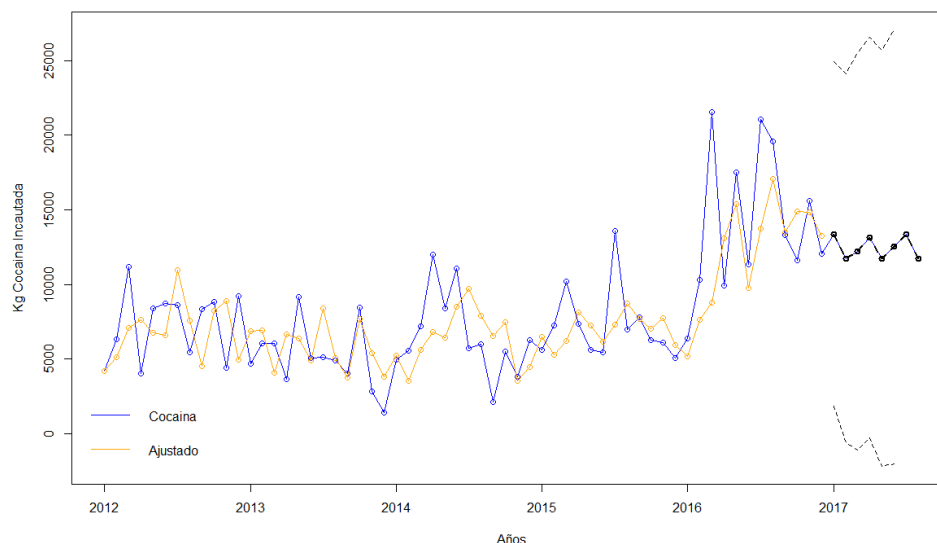
Gráfica 22. Prueba de homocedasticidad del modelo ARIMAX (2.1.4) con 02 regresores



Fuente: Elaboración propia (2017)

Respecto a la prueba de homocedasticidad se concluye que con un nivel de significancia del 5%, hay suficiente evidencia estadística para afirmar que 98% de los residuales son homocedásticos, ósea la varianza es constante para los primeros 20 rezagos. La gráfica 23 muestra la superposición de la serie original (en color azul) y la serie ajustada del modelo del modelo ARIMAX (2.1.4) con 02 Regresores (en color amarillo). En la parte final de la serie original se observan los pronósticos hechos con este modelo para los primeros 6 meses del año 2017.

Gráfica 23. Serie de tiempo original, del modelo ARIMAX (2.1.4) con 02 regresores y pronóstico



Fuente: Elaboración propia (2017)

4.4 PRUEBAS DE LOS MODELOS ARIMAX CON ANÁLISIS DE INTERVENCIÓN

Se procedió luego a realizar las pruebas de modelos ARIMAX con análisis de intervención, incorporando dummies regresoras.

La tabla 7 reúne los resultados de las pruebas de normalidad, de correlación y del BIC para los cinco (5) modelos probados.

Tabla 7. Comparativo de las pruebas de los modelos ARIMAX con análisis de intervención

Análisis Modelos	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
	ARIMAX (7.1.1)	ARIMAX (7.1.1)	ARIMAX (7.1.1)	ARIMAX (7.1.1)	ARIMAX (7.1.1)
	12 Dummies	8 Dummies	7 Dummies	5 Dummies	6 Dummies puntuales y una dummie LS
Criterio Información Bayesiana (BIC)	BIC 434	BIC 425	BIC 423	BIC 416	BIC 380
Normalidad Prueba Jarque Bera Test	1%	6%	44%	46%	85%
Correlacion Prueba Ljung-Box Test	98%	99%	99%	99%	87%

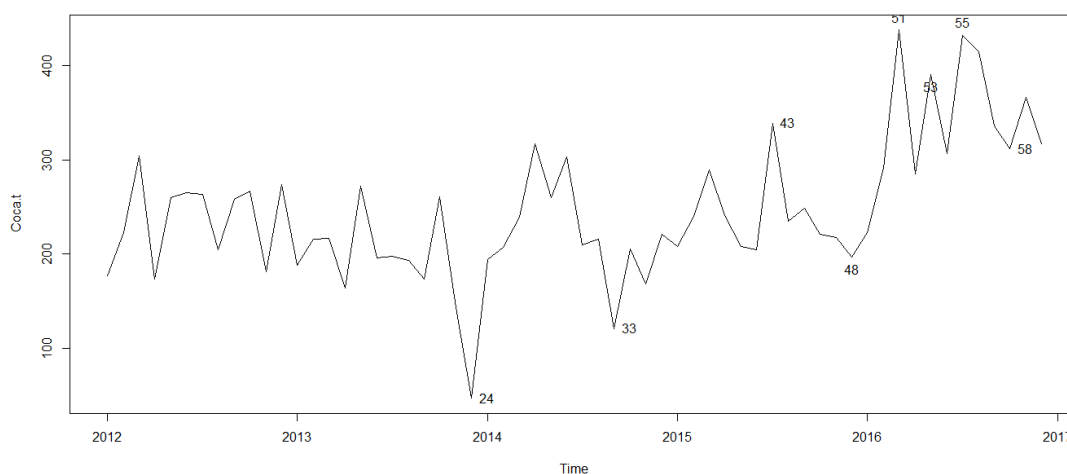
Fuente: Elaboración propia (2017)

Puede apreciarse en la tabla anterior que, en general, estos cinco modelos ARIMAX en cuales se incorporaron análisis de intervención respectivos, presentan un menor BIC que los encontrados hasta ahora en los anteriores modelos.

Con base en la anterior tabla se explican únicamente a continuación los dos (2) modelos que obtuvieron el mejor Criterio de Información Bayesiana (BIC); el modelo ARIMAX (7.1.1) con análisis de intervención de cinco (5) dummies y el modelo ARIMAX (7.1.1) con análisis de intervención de siete (7) dummies, una de las cuales incorpora un cambio de nivel (LS, Level Shift) al final de la serie; este cambio de nivel incorpora los datos de la posición 50 hasta la 60 de la serie transformada, los cuales corresponden al año 2016.

La gráfica 24 muestra la serie transformada y algunos de los puntos atípicos (outliers) observados.

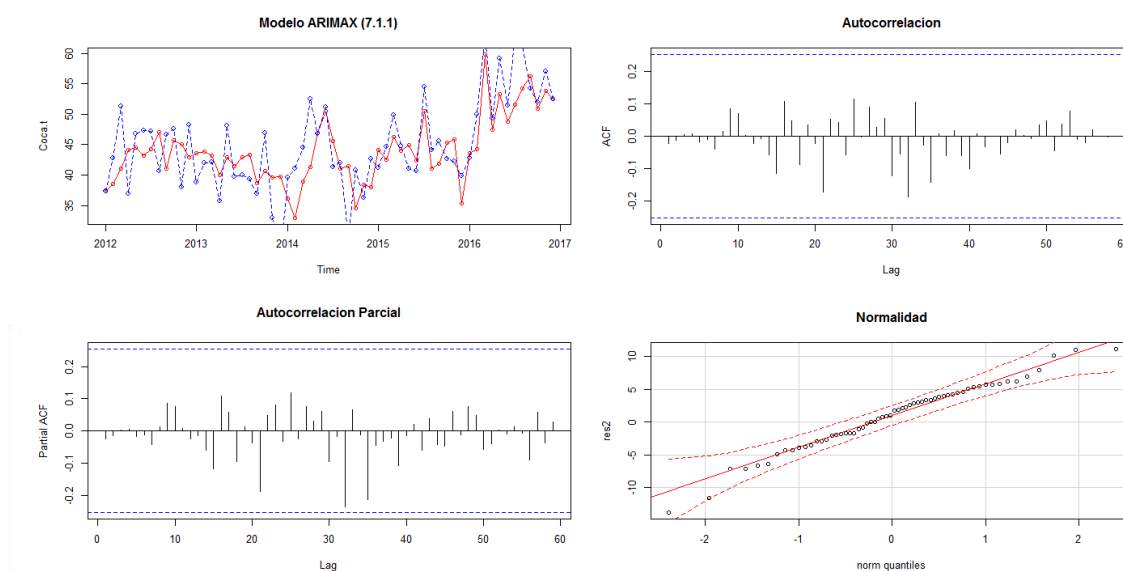
Gráfica 24. Serie de tiempo transformada y algunos valores atípicos



Fuente: Elaboración propia (2017)

4.4.1 Prueba del modelo ARIMAX (7.1.1) con análisis de intervención de 05 dummies. La gráfica múltiple 25 presenta la superposición de la serie original (en color azul) y la serie ajustada del modelo ARIMAX (7.1.1) con 05 dummies (en color rojo), los correlogramas de esta serie ajustada y el gráfico de la prueba de normalidad. Las cinco dummies introducidas al modelo corresponden a I43, I48, I51, I53, I58, que fueron las que resultaron ser más significativas en un modelo previo que incorporó mayor cantidad de dummies.

Gráfica 25. Prueba del modelo ARIMAX (7.1.1) con análisis de intervención de 05 dummies



Fuente: Elaboración propia (2017)

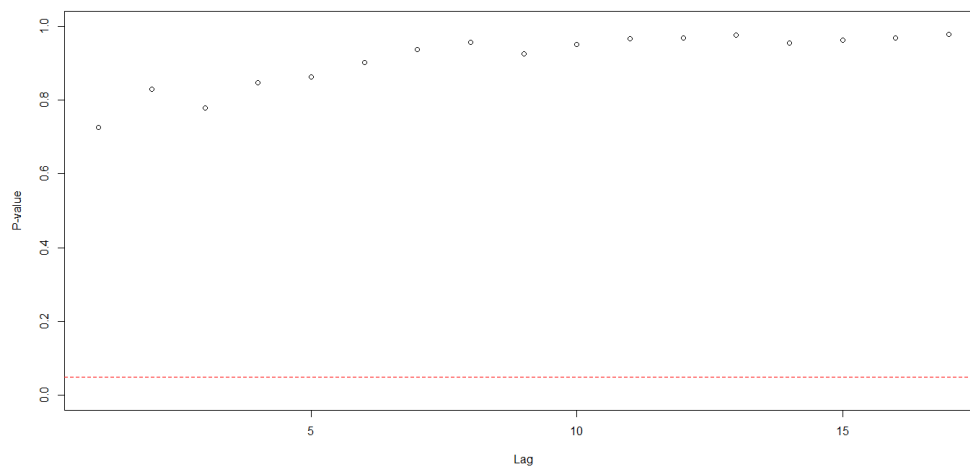
En el gráfico de autocorrelación simple (ACF) se evidencia que todos los rezagos (Lag) se encuentran dentro de la banda de confianza, al igual que en el gráfico de autocorrelación parcial (PACF). En el gráfico de normalidad se muestra que ninguno de los puntos se sale de las bandas de confianza, evidenciando que la serie de tiempo sigue una distribución normal.

Se realiza entonces la prueba de normalidad del modelo a través de la prueba “Jarque Bera Test”, obteniéndose un p-value = 46%, el cual es mayor a $\alpha = 0.05$, lo que quiere decir que se acepta la hipótesis alterna H_1 y se rechaza la hipótesis nula H_0 y se concluye que los residuales siguen un comportamiento normal.

Luego se verifica la correlación de los residuales a través de la prueba “Ljung-Box”, obteniéndose un p-value = 99%, que es mayor a $\alpha = 5\%$, por lo cual se rechaza H_0 (hay correlación) y se acepta H_1 (no hay correlación), lo que quiere decir que los residuales son independientes uno del otro.

La gráfica 26 muestra que los residuales (puntos) se ubican por encima de la banda de confianza, evidenciando la homocedasticidad del modelo ARIMAX (7.1.1) con 05 dummies de Intervención.

Gráfica 26. Prueba de homocedasticidad del modelo ARIMAX (7.1.1) con 05 dummies de intervención

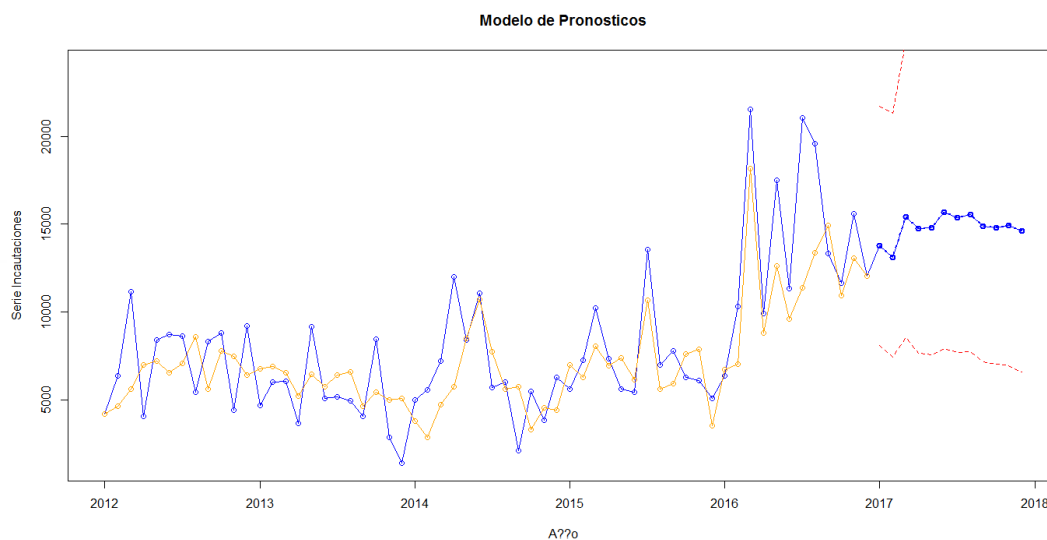


Fuente: Elaboración propia (2017)

Respecto a la prueba de homocedasticidad se concluye que, con un nivel de significancia del 5%, hay suficiente evidencia estadística para afirmar que 94% de los residuales son homocedásticos, o sea, la varianza es constante para los primeros 20 rezagos.

A continuación, la gráfica 27 muestra la superposición de la serie original (en color azul) y la serie ajustada del modelo ARIMAX (7.1.1) con análisis de intervención de 05 dummies (en color amarillo). En la parte final de la serie original se observan los pronósticos hechos con este modelo para los 12 meses del año 2017.

Gráfica 27. Serie de tiempo original y del modelo ARIMAX (7.1.1) con 05 dummies de Intervención y pronóstico

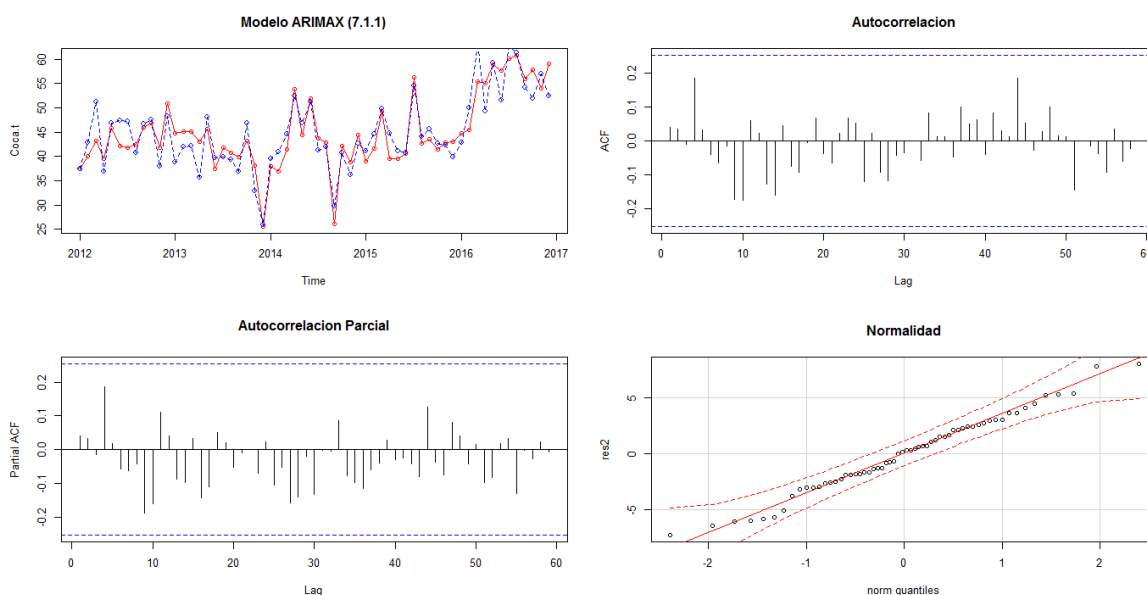


Fuente: Elaboración propia (2017)

4.4.2 Prueba del modelo ARIMAX (7.1.1) con análisis de intervención de 06 dummies tipo puntual y una de cambio de nivel (LS)

La gráfica múltiple 28 presenta la superposición de la serie original (en color azul) y la serie ajustada del modelo ARIMAX (7.1.1) con 06 dummies de tipo puntual y una de cambio de nivel (LS - Level Shift) al final de la serie (en color rojo), los correlogramas de esta serie ajustada y el gráfico de la prueba de normalidad. Los 06 dummies de tipo puntual introducidos al modelo corresponden a I24, I28, I30, I33, I39, I43, y la última dummie (ILS1), de cambio de nivel, comprende los valores 50 hasta el 60 de la serie; todas estas dummies habían resultado ser más significativas en modelos anteriores.

Gráfica 28. Prueba del modelo ARIMAX (7.1.1) con análisis de intervención de 06 dummies de tipo puntual y una de cambio de nivel



Fuente: Elaboración propia (2017)

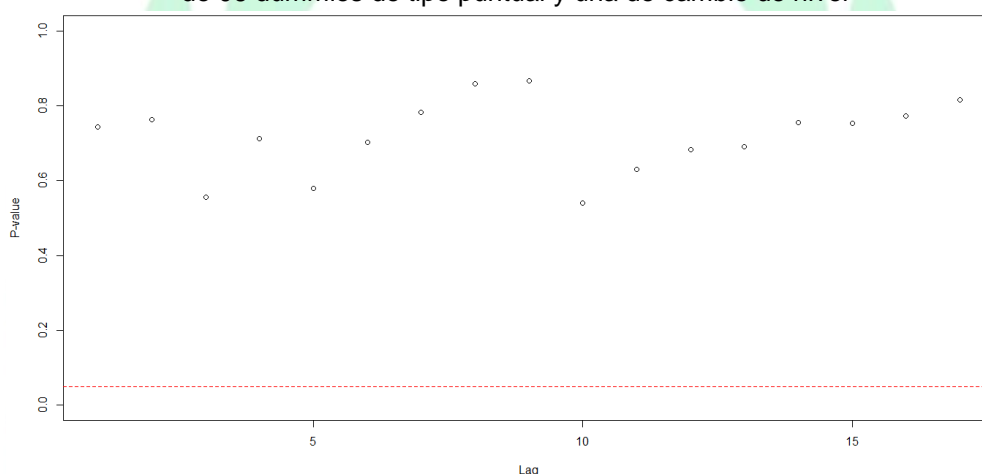
En el gráfico de autocorrelación simple (ACF) se evidencia que todos los rezagos (Lag) se encuentran dentro de las bandas de confianza, al igual que en el gráfico de autocorrelación parcial (PACF). En el gráfico de normalidad se muestra que ninguno de los puntos se sale de las bandas de confianza, evidenciando que la serie de tiempo sigue una distribución normal.

Se realiza entonces la prueba de normalidad del modelo a través de la prueba “Jarque Bera Test”, obteniéndose un p-value = 85%, el cual es mayor a $\alpha = 0.05$, lo que quiere decir que se acepta la hipótesis alterna H_1 y se rechaza la hipótesis nula H_0 , y se concluye que los residuales siguen un comportamiento normal.

Luego se verifica la correlación de los residuales a través de la prueba "Ljung-Box", obteniéndose un p-value = 87%, que es mayor a $\alpha = 0.05$, por lo cual se rechaza H_0 (hay correlación) y se acepta H_1 (no hay correlación), lo que quiere decir que los residuales son independientes uno del otro.

La gráfica 29 muestra que los residuales (puntos) se ubican por encima de la banda de confianza, evidenciando la homocedasticidad del modelo ARIMAX (7.1.1) con intervención de 06 dummies de tipo puntual y una de tipo LS (Level Shift) al final de la serie.

Gráfica 29. Prueba de homocedasticidad del modelo ARIMAX (7.1.1) con análisis de intervención de 06 dummies de tipo puntual y una de cambio de nivel

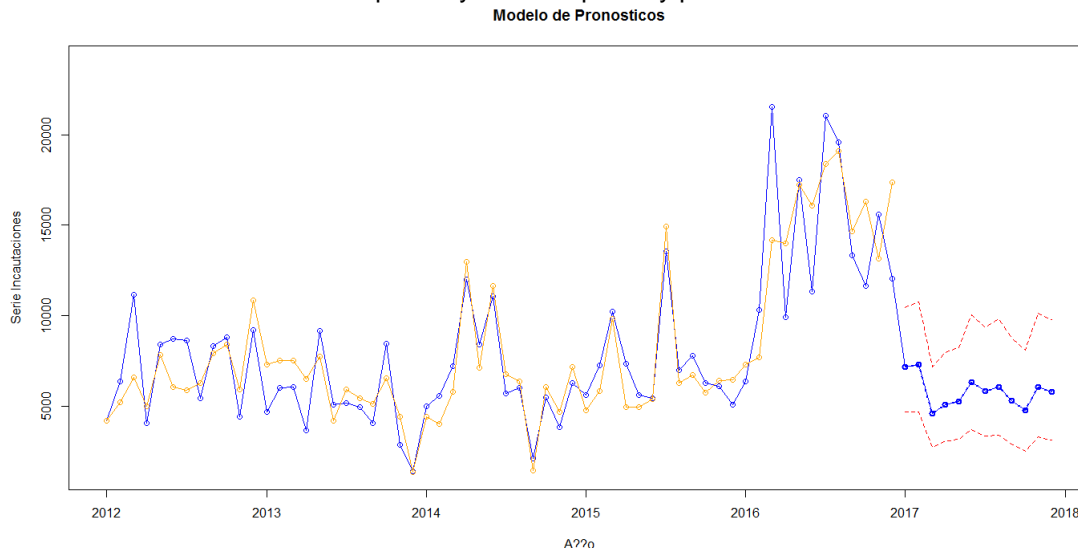


Fuente: Elaboración propia (2017)

Respecto a la prueba de homocedasticidad se concluye que, con un nivel de significancia del 5%, hay suficiente evidencia estadística para afirmar que el 90% de los residuales son homocedásticos, o sea, la varianza es constante para los primeros 20 rezagos.

La gráfica 30 muestra la superposición de la serie original (en color azul) y la serie ajustada del modelo ARIMAX (7.1.1) con intervención de 06 dummies de tipo puntual y una de cambio de nivel (LS) para el último tramo de la serie (en color amarillo). En la parte final de la serie original se observan los pronósticos hechos con este modelo para los 12 meses del año 2017.

Gráfica 30. Serie de tiempo original y del modelo ARIMAX (7.1.1) con intervención de 06 dummies de tipo AO y una de tipo LS y pronóstico



Fuente: Elaboración propia (2017)

4.4.3 Pruebas de capacidad predictiva de los modelos ARIMAX(7.1.1) con intervención de 07 dummies, con 05 dummies y con 07 dummies (una de cambio de nivel - LS). Para realizar las pruebas de capacidad predictiva mediante *back testing* se seleccionaron los tres (3) modelos que cumplieron con los criterios anteriormente mencionados: ARIMAX (7.1.1) con 07 dummies puntuales, el ARIMAX (7.1.1) con 05 dummies puntuales y el ARIMAX (7.1.1) con 06 dummies puntuales y una dummie de cambio de nivel LS.

El modelo de prueba utilizado para obtener el modelo más óptimo fue el MAPE, (Mean absolute percentage error) o Error Porcentual Absoluto Medio. Este permitió comparar entre los tres modelos de pronóstico y obtener el mejor ajuste de las predicciones de los tres (3) modelos en relación con los datos reales de las incautaciones de cocaína realizados por la Armada Nacional desde el mes de enero hasta el mes de mayo de la presente vigencia (2017).

La tabla 8 muestra los resultados obtenidos en las diferencias mensuales de predicción en cada modelo:

Tabla 8. Diferencias mensuales en las predicciones de los tres modelos en relación con los datos reales

MES	VALOR	PRONÓSTICO	ERROR	PRONÓSTICO	ERROR	PRONÓSTICO	ERROR
	REAL	ARIMAX (7.1.1) 07 dummies puntuales		ARIMAX (7.1.1) 05 dummies puntuales		ARIMAX (7.1.1) 06 dummies puntuales y 01 dummie LS	
ene	12.730	10.472	2.258	12.730	0	7.149	5.581
feb	16.645	9.803	6.842	12.113	4.532	7.276	9.369
mar	15.343	10.978	4.365	14.025	1.318	4.568	10.775
abr	16.409	11.623	4.786	14.161	2.248	5.093	11.316
may	10.468	11.689	-1.221	13.955	-3.487	5.276	5.192

Fuente: Elaboración propia (2017)

Obsérvese que los valores de errores de pronóstico más bajos corresponden a los del modelo ARIMAX (7.1.1) con 05 dummies puntuales, lo cual indica que es el más preciso para pronosticar. La tabla 9 muestra el ajuste MAPE de los tres modelos anteriores:

Tabla 9. Mejor ajuste (MAPE) de los modelos ARIMAX (7.1.1) con 07 dummies, el ARIMAX (7.1.1) con 06 dummies y el ARIMAX (7.1.1) 06 dummies puntuales y una dummie LS

MAPE		
ARIMAX (7.1.1) 07 DUMMIES	ARIMAX (7.1.1) 05 DUMMIES	ARIMAX (7.1.1) 07 DUMMIES 01 DUMMIE LS
Yt-Yt ajus	Yt-Yt ajus	Yt-Yt ajus
0,177376277	0	0,561586803
0,411054371	0,272273956	0,437128267
0,284494558	0,085902366	0,297725347
0,291669206	0,136997989	0,310378451
0,116641192	0,333110432	0,504012228
1,281235603	0,828284743	2,110831095
21,35%	13,80%	35,18%

Fuente: Elaboración propia (2017)

Como se puede observar en la tabla 9, los tres (3) modelos ajustan bien, pero al analizar el Error Porcentual Absoluto Medio, el que presenta menor alejamiento frente a los datos reales es el modelo ARIMAX (7.1.1) con 05 dummies; en la fila inferior de la tabla se observan los porcentajes finales del cálculo MAPE, observándose que el modelo ARIMAX (7,1,1) con 05 dummies presenta el menor.

5. CONCLUSIONES Y RECOMENDACIONES

- A partir de los análisis realizados sobre los veintiún (21) modelos estadísticos propuestos modelos ARIMA y ARIMAX, se determinó que los tres (3) modelos que más se ajustaron a la serie de tiempo original son los modelos ARIMAX (7.1.1) con 07 dummies, el ARIMAX (7.1.1) con 05 dummies y el ARIMAX (7.1.1) 07 dummies puntuales y una dummy de cambio de nivel LS. Ya ajustados las dificultades con la distribución normal de estos tres (3), se puede determinar que los residuales son ruido blanco gaussianos.
- Después de realizar las pruebas de capacidad predictiva (back testing) a través del MAPE a los tres (3) modelos que tienen mejor ajuste a la serie de tiempo original, ARIMAX (7.1.1) con 07 dummies, el ARIMAX (7.1.1) con 05 dummies y el ARIMAX (7.1.1) con 07 dummies puntuales y un dummy de cambio de nivel LS, se estableció que el modelo más conveniente es ARIMAX (7.1.1) con 05 dummies, ya que este presenta menores variaciones frente a los datos reales de comparación.
- De acuerdo con lo anterior se proyecta la meta operacional de incautaciones de clorhidrato de cocaína para la Armada Nacional en la vigencia 2017, tomando como referencia el modelo que presentó menos variación frente a los datos reales - ARIMAX (7.1.1) con 5 dummies; se incautarán aproximadamente **170** toneladas de clorhidrato de cocaína este año.
- La disminución de las incautaciones de clorhidrato de cocaína realizadas por la Armada Nacional para la vigencia 2017 en comparación al año anterior, se explican debido al incremento de erradicaciones de cultivos de coca mediante la modalidad de aspersión terrestre. Es de aclarar que en los años 2015 y 2016 se suspendieron las erradicaciones aéreas de cultivos de coca mediante aspersión de glifosato, lo cual puede explicar en gran parte el incremento de la producción de cocaína en esos años; sin embargo, para contrarrestar ese impacto, este año en Colombia se han aumentado las acciones de erradicación por aspersión terrestre (Dirección de Operaciones Navales, 2017).
- Es recomendable desarrollar modelos multivariados que permitan integrar todas las variables asociadas al narcotráfico y que son de interés para la Armada Nacional.
- Es recomendable para la Armada Nacional desarrollar modelos de tipo ARIMA y ARIMAX para los demás fenómenos criminales de su interés, diferentes al narcotráfico.

REFERENCIAS

Cryer, J.; Chan, K. (2009). *Time Series Analysis With Applications in R*. Segunda edición. USA: Springer International Edition.

De Arce, R.; Mahía, R. (s.f.). *Modelos ARIMA*. [en línea]. Programa Cititus. Universidad Autónoma de Madrid. Extraído de:
https://www.uam.es/personal_pdi/economicas/anadelsur/pdf/Box-Jenkins.PDF

Detting, M. (2015). *Applied Time Series Analysis*. Institute for Data Analysis and Process Design. Zurich University of Applied Sciences. Extraído de:
https://stat.ethz.ch/education/semesters/ss2015/atsa/ATSA_Scriptum_v1_SS15.pdf

García, Márquez R., Almirante, Comandante Armada Nacional de Colombia. (2011). Plan Estratégico Naval 2011-2014., Armada Nacional de Colombia, Bogotá, Colombia.

Gujarati, D. (2004). *Econometría*. Cuarta edición, México D.F.: Mc Graw-Hill Interamericana Editores S.A. Extraído de:
https://www.dropbox.com/sh/Ineqff1ohmy2dhi/AAC0pRtWB5Rhmc5t_J4t-GkSa/Econometria%20-%20Damodar%20N.%20Gujarati%20%284ta%20Edicion%29.pdf

Gujarati, D. (2010). *Econometría*. Quinta edición, México D.F.: Mc Graw-Hill Interamericana Editores S.A. Extraído de:
https://scalleruizunp.files.wordpress.com/2015/04/econometria_-_damodar_n-_gujarati.pdf

Hernández, J. (2009). *Análisis de Series Temporales Económicas II*. Segunda Edición. Madrid, España: ESIC Editorial

Mascareñas, J. (2013). *Procesos estocásticos: introducción*. Monografías de Juan Mascareñas sobre Finanzas Corporativas. N° 27. Universidad Complutense de Madrid. Extraído de: <http://pendientedemigracion.ucm.es/info/jmas/mon/27.pdf>

Mauricio, J. (2007). *Introducción al Análisis de Series Temporales*. [en línea]. Universidad Complutense de Madrid. Extraído de:
<https://www.ucm.es/data/cont/docs/518-2013-11-11-JAM-IASST-Libro.pdf>

Observatorio de Drogas de Colombia. (2016). Mercado. Bogotá, Colombia: Observatorio de Drogas de Colombia ODC. Extraído de:
<http://www.odc.gov.co/problematika-drogas/oferta-drogas/mercado>

Oficina de las Naciones Unidas contra la Droga y el Delito – UNODC. (2014). Colombia - Monitoreo de Cultivos de Coca 2013. Extraído de:
<http://www.odc.gov.co/Portals/1/publicaciones/pdf/oferta/simci/simci04012014-precios.pdf>

Pinzón Bueno J., Ministro de la Defensa Nacional. (2014) Libro blanco de las estadísticas del sector seguridad y defensa, Ministerio de Defensa Nacional, Bogotá, Colombia.

Rodríguez, Barragán J., Comandante General de las Fuerzas Militares de Colombia, (2014), Directiva Transitoria No 168. Evaluación Plan de Guerra “Espada de Honor” IV., Comando General de las Fuerzas Militares, Bogotá, Colombia.

Sánchez, F., Díaz, A., & Formisano, M. (2003). *Conflicto, violencia y actividad criminal en Colombia: un análisis espacial*. Documento CEDE 2003-05. Centro de Estudios sobre Desarrollo Económico (CEDE), Universidad de los Andes. Recuperado de: <https://core.ac.uk/download/pdf/6517007.pdf>

Santamaría Gaitán L, Almirante, Comandante Armada Nacional de Colombia. (2015). Plan de Campaña “Hidra” Armada Nacional de Colombia, Bogotá, Colombia.

Shumway, R.; Stoffer, D. (2006). *Time Series Analysis and Its Applications With R Examples*. New York, USA: Springer Science+Business Media, LLC

